THE MASS SPECTROMETRY OF SELECTED PROTEINS AND PEPTIDES

Alan L. Millar

Submitted to the University of Warwick in fulfilment of the requirements for the award of Doctor of Philosophy

University of Warwick
Departments of Chemistry and Biological Sciences

September 1998

Table of Contents

Contents

Acknowledgements

I would like to thank my academic supervisor Professor Keith R. Jennings whose tuition and guidance throughout my time at the University of Warwick has been exemplary.

I would also like to express my gratitude to Professor Howard Dalton for welcoming me into his research group and his enthusiastic interest in my studies.

I would like to offer my sincere thanks to the academics and researchers it has been my pleasure to collaborate with during the course of this work they include; Dr Armelle Buzy, Dr Steve Gallagher, Miss Sue Slade, Dr David Griffiths, Dr Nicholas Jackson and Prof. Nigel Dimmock.

I acknowledge EPSRC and Micromass UK for CASE award funding and the British Mass Spectrometry Society (BMSS) and the American Study and Student Exchange Committee (ASSEC) for financial assistance towards the cost of travelling to conferences.

Dr Brian Green of Micromass UK has been the source of useful discussions and encouragement.

I have made many good friends during my time at Warwick both in the departments of Chemistry and Biological Sciences. I thank you all and wish each and every one of you long and happy careers.

Finally, I thank my family for all support they have given me throughout my studies. My parents Alan and Patricia are, and will always be, my inspiration.

Declaration

I hereby declare that this thesis contains my own work. Acknowledgements of relevant information from other studies is made in the text.

Papers Presented at the Following Conferences:

22th Annual Meeting of BMSS, University of Swansea, September 1996

Poster presentation: 'Characterisation of the Alkene Monooxygenase by ESI-MS'

3rd Portuguese International Mass Spectrometry Conference, Lisbon, September 1997

Oral presentation: 'Characterisation of the Ceroid Lipofuscinosis Protein by ESI- MS

Poster presentation: 'Characterisation of the Alkene Monooxygenase by ESI-MS'

The Jennings Symposium, University of Warwick, December 1997

Poster presentation: 'Epitope identification in HIV-1 peptides recognised by a Micro-Antibody by means of electrospray ionisation mass spectrometry'

47th Annual Conference of the American Society of Mass Spectrometry, Orlando, USA, June 1998.

Poster Presentation: An ESI-MS study of the Binding of a 17 Amino Acid Micro-Antibody to V3 peptides of the HIV-1 Envelope Protein.

Poster Presentation: ESI-MS Study of the β-subunit of the Soluble Methane Monooxygenase.

23th Annual Meeting of the British Mass Spectrometry Society, University of Warwick, September 1998.

Oral Presentation: 'Rapid Analysis of Epitope-Paratope Interactions Between HIV-1 and a 17 amino acid Micro-Antibody by means of ESI-MS'

Publications

'Rapid Analysis of Epitope-Paratope Interactions Between HIV-1 and a 17 Amino Acid Micro-Antibody by Electrospray Ionisation Mass Spectrometry' A. L. Millar, N. A. C. Jackson, H. Dalton, K. R. Jennings, M. Levi, B. Wahren, N. J. Dimmock, Accepted for Publication, *Eur. J. Biochem.*

'The Hydroxylase Component of Soluble Methane Monooxygenase from *Methyloccus Capsulatus* (Bath) Exists in Several Forms as shown by Electrospray Ionisation Mass Spectrometry' A. Buzy, A. L. Millar, V. Legros, P. C. Wilkins, H. Dalton, K. R. Jennings, *Eur. J. Biochem.*, **254,** 602 (1998).

'Electrospray Ionisation Mass Spectrometry Studies of the Interaction of N'N - dicyclohexylcarbodiimide with Ceroid Lipofuscinosis Protein (Subunit C)' D. Griffiths, A. Buzy, K. R. Jennings, A. L. Millar, E. M. Ryan, D. Palmer, *Eur. Mass Spectrom.*, **3,** 81 (1997).

List of Abbreviations

a	Mathieu Stability Parameter
a.c.	Alternating Current
AIDS	Acquired Immunodeficency Syndrome
AMO	Alkene Monooxygenase
APcI	Atmospheric Pressure Chemical Ionisation
ATP	Adenosine Triphosphate
CDR	Complementary Determining Region
CF-FAB	Continuous Flow FAB
CI	Chemical Ionisation
CID	Collision Induced Decomposition
CLP	Ceroid Lipofuscinosis Protein
CNBr	Cyanogen Bromide
Da	Dalton
DCCD	Dicyclohexylcarbodiimide
d.c.	Direct Current
DNA	Deoxyribonucleic Acid
DTT	Dithiothritol
e	Elementary Unit of Charge
E_{COM}	Centre of Mass Kinetic Energy
EI	Electron Impact

E_{lab}	Initial Kinetic Energy of the Ion
ELISA	Enzyme Linked Immunsorbent Assay
ESI	Electrospray Ionisation
FAD	Flavin Adenine Dinucleotide
eV	Electron Volts
FAB	Fast Atom Bombardment
FD	Field Desorption
FI	Field Ionisation
FPLC	Fast Protein Liquid Chromatography
FT-ICR	Fouier Transform Ion Cyclotron Resonance
gp	Glyco-Protein
h	Rf only Hexapole
HIV-1	Human Immunodeficency Virus Type-1
HPLC	High Performance Liquid Chromatography
H.V.	High Voltage
k	Kilo
LSIMS	Liquid Secondary Ion Mass Spectrometry
m	Mass of an Ion
Mab	Monoclonal Antibody
MALDI	Matrix Assisted Laser Desorption Ionisation
MCA	Multichannel Acquisation
MCP	Microchannel Plate Detector

m_i	Mass of a Precursor Ion
Micro-Ab	Micro-Antibody
m_p	Mass of the Projectile Ion
m_{pi}	Impact Portion of the Projectile Ion
Mol	mole
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MS1	First Mass Spectrometer of a tandem Instrument
MS2	Second Mass Spectrometer of a tandem Instrument
MS^3	CID Experiment which Produces Second Generation Fragment Ions
m_t	Mass of the Target Gas
m/z	Mass / Charge
N	Number of Atoms in an Ion
NADH	Nicotinamide Adenine Dinucleotide (reduced)
NMR	Nuclear Magnetic Resonance
pmol	Picomole
ppm	Parts per million
PTM	Post Translational Modification
RF	Radio Frequency
RMM	Relative Molecular Mass
q	Mathieu Stability Parameter
Q	Quadrupole

QET	Quasi-Equilibrium Theory
QIT	Quadrupole Ion Trap
Q1	First Quadrupole Analyser in a tandem Quadrupole Instrument
Q2	Second Quadrupole Analyser in a tandem Quadrupole Instrument
R	Resolution Required to Separate Two Peaks of Mass m and Δm
R.F.	Radio Frequency
2ro	Spacing Between Opposite Rods of a Quadrupole
t	Time
t_c	Interaction Time of the Ion and target Species during CID
TCLA	T-Cell Lab Adapted
TFA	Trifluoroacetic Acid
TIC	Total Ion Current
TML	Trimethyl-Lysine
TOF	Time of Flight
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
sMMO	Soluble Methane Monooxygenase
U	Direct Current Voltage Applied to a Quadrupole
UV	Ultra Violet
V	Volts
v	Velocity
Vr	Retarding Potential

ω	Frequency Measured in Radians of the R.F. Voltage Applied to a Quadrupole
z	Number of Elementary Charges on an Ion

Abstract

The multi-subunit alkene monooxygenase was studied by means of electrospray ionisation mass spectrometry (ESI-MS). The measured relative molecular mass (RMM) values for the epoxygenase and reductase were in good agreement the predicted RMM values calculated by means of gene sequencing. Cleavage sites were identified by means of ESI-MS at the N- and C- terminus of the coupling protein. Cleavage at the C-terminus inactivated the protein but was shown to be prevented by the addition of a protease inhibitor, benzamidine.

The ceroid lipofuscinosis protein (CLP) was found to contain a +42.2 Da modification. Low energy tandem mass spectrometry confirmed 80% of the primary sequence of the CLP, indicating that the discrepancy in RMM is located on amino acid residues 30-46. Bovine heart mitochondrial subunit c was extracted as a control and was also found to contain the +42.2 Da modification. A study of the interaction of CLP with its diagnostic inhibitor DCCD was performed. This indicated that the primary product of the reaction was an acetyl-CLP complex and not the expected CLP-DCCD complex.

The β-subunit of the soluble methane monooxygenase (sMMO) was investigated by means of ESI-MS to locate and identify a +126.3 Da discrepancy in RMM. A tryptic digestion of the protein indicated the discrepancy to be located in a small digest peptide consisting of amino acid residues 141-146. Low energy MS/MS was used to correct the primary sequence of this peptide. The revised sequence was supported by the close similarity between the DNA nucleotide sequence for the originally suggested amino acid residue sequence and the revised sequence.

The binding between a 17 amino acid micro-antibody (Micro-Ab) and two peptides which represent the V3 region of the gp 120 external glycoprotein of the HIV-1 envelope protein were studied by means of ESI-MS. It was found that the interaction between the Micro-Ab and the two V3 peptides could be maintained and thus observed directly by ESI-MS. A rapid method of epitope identification was devised whereby the antigen was pre-digested before incubation with the Micro-Ab. The data identified a similar series of amino acid residues in both V3 peptides as the epitope with which the Micro-Ab interacts.

CHAPTER 1

INTRODUCTION

1.1 Introduction to mass spectrometry

Mass spectrometry (MS) is an analytical technique which one can use to determine accurately the relative molecular mass of an ionised sample. Mass spectrometry offers both high sensitivity and selectivity and thus finds useful applications in many areas of research. Furthermore, tandem mass spectrometry (MS/MS) can be used to ascertain structural information about a sample. This can be useful in identifying compounds in complex mixtures and determining structures of unknown materials.

There are several different types of mass spectrometer available, each with it's own advantages and disadvantages. The choice of mass spectrometer is highly dependent upon the application. The different types of instrument are normally classified by the mass analyser they employ. All mass spectrometers are composed of four distinct sections; ionisation source, mass analyser, ion detection and data system as represented in Figure 1.1.

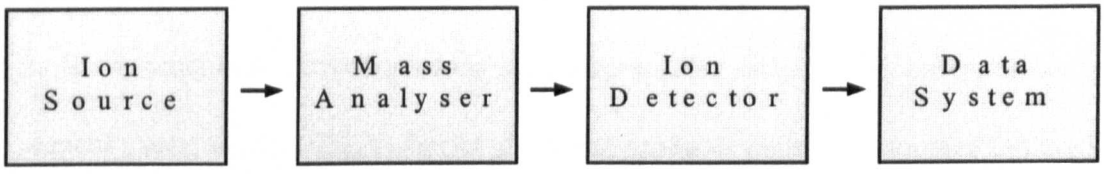


Figure 1.1. The four sections of a mass spectrometer.

Although the origins of mass spectrometry can be traced back to the beginning of the century with the work of J.J. Thomson [1], it was not until the 1940s that the first commercial instruments were available. These instruments were developed for analysis of hydrocarbons in the oil industry. At that time the only ionisation source available was electron impact (EI) developed by Nier [2] after its conception by Dempster [3].

EI involves bombarding a neutral gaseous sample with a beam of energetic electrons. The energetic electrons interact with the sample by passing close to it or even through the molecule to eject an electron from the neutral molecule. Thus two electrons leave the reaction region, leaving a positively charged molecular ion. At 70 electron volts (eV) these interactions can leave the molecular ion with a range of internal energies causing extensive fragmentation to occur. This can be very desirable for gaining structural information about the sample molecule, although it can result in a low intensity parent ion peak making relative molecular mass determination difficult. EI proved to be very reproducible and popular and still finds widespread use today.

Chemical ionisation (CI), developed in 1966 [4], generates ions by means of ion-molecule reactions between gaseous sample molecules and, in many cases, protonated reagent gases, such as CH_5^+. As the technique produces mainly protonated molecular ions of the type (MH^+) and promotes little fragmentation it is deemed to be a milder ionisation technique than EI.

Both EI and CI suffer from a major weakness: they require the sample to be in the gas phase and thus cannot accommodate large involatile molecules or thermally sensitive compounds. This restriction means that large polar compounds, such as proteins and peptides, are usually unsuitable for analysis by EI / CI due to their low volatility and frequent large size.

Ionisation of thermally labile, involatile molecules was made possible with the introduction of field ionisation [5] and field desorption [6] (FI / FD). Using these techniques the analyte is ionised by intense local electric fields. It was, however, the introduction of fast atom bombardment mass spectrometry (FAB-MS) [7] in 1981 that started what has subsequently become a rapidly expanding area of research: biological mass spectrometry.

FAB employs a high energy beam of fast atoms to bombard a liquid target surface to produce continuous desorption of ions which are characteristic of the liquid. When a sample is dissolved in a solvent of low volatility, both positive and negative molecular ions or quasimolecular sample ions are produced. FAB is performed by mixing a sample solution of approximately 1μL with approximately 1-2 μL of matrix and placing it on a target, which is usually the tip of a direct insertion probe. Upon insertion into the mass spectrometer, the sample is bombarded with fast atoms.

It was later found that a primary beam of fast ions could utilised instead of atoms, this

technique was termed liquid secondary ion mass spectrometry (LSIMS) [8]. A development of standard or static FAB was continuous flow FAB (CF-FAB) [9] which makes use of a flowing carrier solution which continuously refreshes the target surface. This has the advantage of allowing the sample to be introduced in a liquid form for mass spectrometric analysis and of significantly reducing the amount of matrix required in the target mixture.

CF-FAB also facilitates the union of high performance liquid chromatography- mass spectrometry (HPLC-MS) [10]. In general FAB and LSIMS gives excellent results up to approximately 2000 mass / charge (m/z), but above this value the abundance of the molecular ions usually tends to decrease until in the region of 5000-6000 m/z the ions become difficult or impossible to observe. This means that peptides and digest peptides of proteins can be analysed, but larger intact proteins are not suited to this method of ionisation.

Unquestionably the two ionisation techniques which have had the largest impact on biological mass spectrometry are electrospray ionisation (ESI) and matrix assisted laser desportion ionisation (MALDI). ESI has been used throughout this work and is discussed in greater detail in section 2.2.

During MALDI, first described in 1987 [11,12], the sample is mixed with a large molar excess of matrix and placed on a target, this is in contrast to the earlier technique of laser desorption (LD) [13] which did not use a matrix. Typically the analyte is

dissolved in a matrix which absorbs strongly at the wavelength of the laser, e.g. 337 nm for a nitrogen laser. A drop of this solution is then allowed to evaporate slowly on a target. The target is then placed into the ion source of a suitable, typically time of flight (TOF) because of the pulsed ion source, mass spectrometer and irradiated with a laser. Provided the matrix absorbs at the particular wavelength of the laser, energy is transferred to the matrix and by one of several mechanisms ions are transferred to the gas phase.

Sample preparation is of particular importance in MALDI-MS if high quality data are desired. Several different methods have been described for the production of MALDI sample targets such as the dried droplet method [11], the thin polycrystalline film [13] and electrospraying the sample solution onto the target [14]. Ionisation of the sample by a laser differs from FAB/LSIMS as the energetic primary beam is made up of photons which have no significant momentum. MALDI has found widespread use in the study of proteins and peptides and in conjunction with ESI permits the mass spectrometric analysis of a large range of biological samples.

The second section of a mass spectrometer is the mass analyser. Its function is to separate the ions, formed in the ion source, according to their m/z ratio. At present there are five types of commonly encountered mass analysers; magnetic sectors, quadrupole mass filters, time of flight (TOF), quadrupole ion trap (QIT) and fourier-transform ion cyclotron resonance (FT-ICR) instruments. The two types of mass

analyser used during this work were a quadrupole mass filter and a TOF mass analyser, hence these are discussed in greater detail in sections 2.3(a) and 2.4(a) respectively. The three commonly encountered mass analysers not utilised in this work have found interesting applications in the study of proteins and peptides.

Magnetic sector instruments coupled with FAB and ESI have been used to good effect in this area of research due to their excellent resolution and mass accuracy capabilities [15,16,17]. Increasingly however these types of instrument are being superseded by smaller, simpler types of mass spectrometer for biological mass spectrometric applications. This is in part because magnetic sector instruments are large, expensive, complex in operation and pose considerable technical difficulties when coupling atmospheric pressure ionisation techniques to the large accelerating potentials they employ within the ion source region. Another important factor in their demise is that instruments such as MALDI-TOF, ESI- quadrupole, ESI-TOF and QIT are constantly improving in performance to a point where, although they may not be a match for the specifications of some sector instruments, they are often able to answer the biological problem under investigation.

The QIT was first described in 1953 [18], when it was used primarily as a storage device for spectroscopic analysis. It was not until 1984 with the introduction of a new rapid method of scanning mass spectra [19] that interest in this type of mass analyser started to occur. The QIT is a small simple instrument, although the theory of ion motion within the trap itself is very complex. The commercial QIT instruments available tend

to be coupled to ESI [20,21]. In these instruments octapole ion guides are used to transfer the ions from the atmospheric pressure of the ESI source to the very low pressures at which the trap is required to operate. Typically a commercial ion trap instrument currently available has a resolution of approximately 8000 at m/z 2000. The mass range is also very good with proteins up to 100 kDa being easily investigated. Because the QIT is an ion storage device, unlike most conventional instruments, it is particularly suitable for MS^N experiments in time rather than space, with the use of minimum sample[22].

FT-ICR-MS offers the ultimate performance both in terms of resolution and mass accuracy. First described in the mid 1970s [23] FT-ICR-MS has found most success in the analysis of bio-molecules when used in combination with ESI.

FT-ICR-MS can produce ultra-high resolution spectra; for example McLafferty and co-workers [24] reported a mass resolution in excess of 6×10^4 for the 17^+ charge state of cytochrome c at 773 m/z. Later this value was increased when the same workers reported a resolution of 2×10^6 for the 10+ charge state of ubiquitin at 857 m/z [25]. Like the QIT the FT-ICR-MS is capable of MS^N. This makes for a very powerful analytical technique when combined with the ultra high resolution of these instruments which makes it feasible to assign high charge state fragment ions. When used with ESI FT-ICR-MS has an extremely high mass range. Smith and co workers have demonstrated that a molecule, a coliphage T4 DNA, with a molecular weight of 1.1×10^8 can be detected by ESI/FTICR [26].

The mass accuracy achievable by FTICR is also very high. In 1991 McLafferty reported a mass accuracy of 1 parts per million (ppm) for isotopically resolved myoglobin [27]. As with all mass spectrometers there are disadvantages with the FTICR instrument; these include the extremely high cost, complex operation and the limitations imposed by the extremely low pressures at which the ICR cell operates.

For all types of mass analyser except the FT-ICR instrument, ions are transported towards a detector system which measures their relative abundance, ion detection is discussed in detail in sections 2.3 (b) and 2.4 (b). In the FT-ICR instrument, ions are detected by the absorption of radio frequency (R.F.) power and so, uniquely, they are not destroyed at the moment of detection. The final section of the mass spectrometer, the data system, is then used to process the data to produce the mass spectrum.

1.2 Biological Mass Spectrometry

In the last 20 years biological mass spectrometry has emerged as an increasingly important area of research. This is due, most recently, to the development of ESI and MADLI which has enabled the ionisation of both peptides and large proteins. Peptides and proteins consist of a variable number of amino acid residues linked together by peptide bonds. The difference between peptides and proteins is one of size; peptide refers to a molecule composed of a small number of amino acid residues, whilst proteins are much larger. The most important feature of proteins is that their larger size means

that they posses well defined three-dimensional structures. For example, haemoglobin, which transports oxygen from the lungs to the rest of the body has a precise shape which allows oxygen to bind reversibly [28]. In contrast peptides do not possess a well defined three-dimensional structure.

There are twenty common α–amino acids [29]. They all have the same generalized structure shown in Figure 1.2, the only difference between them is the nature of the R-group which specifies the amino acid.

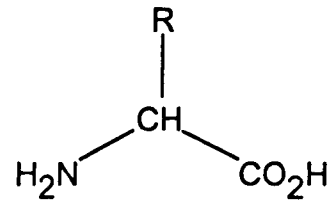


Figure 1.2. Generalised amino acid structure, where R characterises the amino acid type.

Protein structure can be classified as primary, secondary, tertiary and quaternary [30]. The primary structure describes the order of the amino acid residues of which the protein

is composed, whilst the secondary structure refers to regular features of the protein such as α-helices and β-sheets. Tertiary structure describes the three-dimensional shape of the protein and the quaternary structure refers to the arrangement of proteins within multi-subunit proteins.

1.2 (a) Relative molecular mass determination / primary sequence determination

There are many questions about a protein which can be answered by the accurate determination of its relative molecule mass (RMM).

Frequently the primary structure of a protein has been deduced from the base sequence of the gene coding for the protein. The predicted amino acid residue sequence can thus be imported into a protein data base, from which a predicted RMM value for the protein can be obtained. The correctness of the sequence can then be checked in the first instance by comparing the predicted and measured RMM values of the protein [31]. If the two values agree then this is a good indication that the deduced amino acid residue sequence is correct.

If the two values disagree it is an indication of an error in the DNA deduced amino acid residue sequence. Such a mismatch can be caused either by the misidentification of a nucleotide base, which may change the base triplet to the colon for a different amino acid residue, or by missing a base (deletion) or by inserting a fictitious base (insertion) during the interpretation of the sequencing gel [32]. In this situation or when greater mass accuracy is required to differentiate mutations that produce a 1 Da difference (i.e.

Asp and Asn, Leu/Ile and Asn, Glu and Gln), the peptide mapping approach can be used.

Following this methodology the protein is digested by a highly specific proteolytic reagent. Trypsin is a common choice, it hydrolyses peptides and proteins on the carboxyl side of arginine and lysine amino acid residues [33]. These residues occur with moderate frequency in most proteins, so that trypsin digestion usually provides good sequence coverage without resulting in too many digest peptides which would hinder their assignment. If disulfide bonds are present in the protein or peptide to be digested these typically have to reduced and alkylated prior to digestion; failure to do this often results in reduced sequence coverage.

The measured digest peptide RMM values are then matched against theoretical values calculated by means of a protein data base. It is an indication of a possible error in the DNA sequence if one or more of the measured RMM values do not correspond to any of the predicted values. This step has the double advantage of allowing a more accurate determination of the error in RMM and more a more specific location of the error to be determined. Occasionally knowing the RMM value of the discrepancy together with knowledge of the primary sequence of the digest peptide that contains the discrepancy will be sufficient to correct the DNA sequence. This is due to the fact that amino acid residue mutations have defined differences in RMM associated with them. It is, however, not uncommon for multiple mutations to occur [32]. In this situation the total discrepancy in RMM of the protein/digest peptide will be the algebraic sum of the

individual mutations; if this value corresponds to that of a single mutation a mis-correction could easily result.

A more reliable way to assign these mutations is to submit the digest peptide that contains the modification for MS/MS analysis. Tryptic digest peptides are well suited for MS/MS analysis as they usually fall within a suitable mass range for low energy MS/MS analysis (< 2000 Da); also they are often doubly charged and hence fragment to produce singly charged product ions which are relatively simple to assign [34]. MS/MS and collision induced decomposition (CID) are discussed in greater detail in sections 2.5 and 2.5 (b) respectively.

The nomenclature for peptide fragmentation along the backbone of the molecule is such that fragment ions with the charge residing on the N-terminus of the peptide are denoted a_n, b_n and c_n product ions; fragment ions with the charge residing on the C-terminus are identified as x_n, y_n and z_n ions, where n designates the amino acid residue number [35] as illustrated in Figure 1.3.

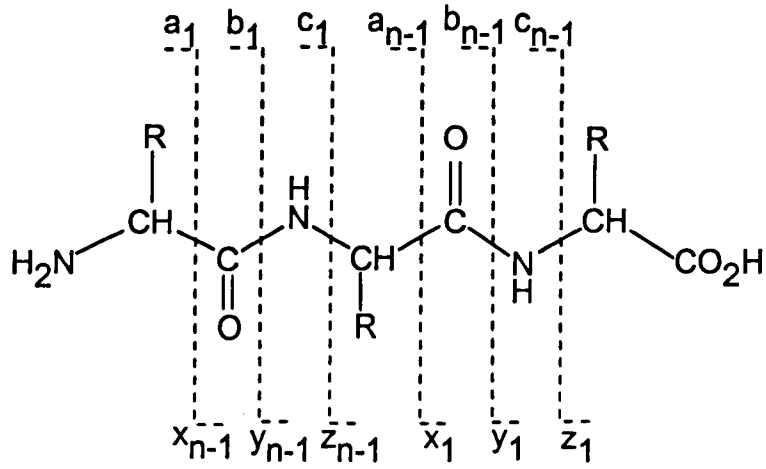


Figure 1.3. Nomenclature describing the fragmentation of peptides.

Identifying the sites of post translational modifications (PTMs) is also another important application of MS and MS/MS [36,37,38]. Common examples of such PTMs include the removal of amino acid residues at the N or C termini after processing, acylation, glycosylation, phosphorylation or sulfonation of certain amino acid residues. Again accurate RMM determination may provide sufficient information to identify the PTM. A frequent PTM is the removal of the N-terminal methionine at some stage after synthesis which is identified by a measured RMM 131 Da lower than expected, as encountered in chapter 3. The peptide mapping approach is also applicable to location and identification of PTM.

Another option to locate certain PTMs such as phosphorylation [39] or glycosylation [40] is to use parent ion scanning. In this experiment, typically performed on a tandem

quadrupole instrument, the first mass analyser (MS1) scans for a parent ion that fragments to produce a product ion characteristic to the modified amino acid residue. For example phosphopeptides containing phosphoserine, phosphothreonine or phosphotyrosine yield fragment ions of m/z 63 (PO2-) and m/z 79 (PO3-); glycoslated peptides are identified by a diagnostic m/z 204 fragment ion.
This product ion is mass analysed in the second mass analyser (MS2) and subsequently detected. This method would allow the identification of a modified peptide from, for example, a digest peptide mixture which could then be submitted for MS/MS analysis [41].

An accurate RMM determination can also confirm the presence or absence of a predicted disulfide bond in a peptide or protein if the amino acid sequence deduced from the use of DNA is available. This is possible as for each disulfide bond present the mass of two hydrogen atoms is subtracted from the peptide/protein RMM. An example of the confirmed presence of a predicted disulfide bond is that of the micro-antibody in chapter 6. If prior sequence information is not available, methods for locating disulfide bonds in peptides and protein exist. These experiments normally involve reduction and alkylation of the covalent disulfide bonds in the protein [36,42].

The analysis of hydrophobic membrane proteins presents a particular problem due to their insolubility in the solvents typically used for ESI and MALDI. Schindler *et al* [43] reported that such proteins can be dissolved in hexafluoroisopropanol or neat formic acid and electrosprayed in a chloroform / methanol / water solution. Other workers have

examined a number of detergents which are suitable for ESI-MS analysis and may help to solubilize membrane proteins for analysis. The hydrophobic subunit c protein described in chapter 4 has successfully been dissolved and sprayed in a chloroform / methanol /water solvent in which the constituents are in the ratio 4/4/1.

More traditional biochemical methods for the determination of RMM and primary sequence of proteins and peptides include sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [44] and Edman degradation [45] respectively.

SDS- PAGE analysis is a widely used analytical technique for the separation and characterisation of complex mixtures of proteins and for the estimation of the relative molecular masses of proteins. Protein mixtures analysed by this method are separated largely on the basis of electrophoresis, the movement of charged particles under the influence of an electric field, under denaturing conditions.

The protein mixture is dissolved in a solution of SDS; this has the effect of disrupting almost all of the non-covalent interactions of the proteins. Disulfide bonds are also reduced by the addition of mercaptoethanol or dithiothritol (DTT). SDS anions bind to the protein in a ratio of approximately one SDS for every two amino acid residues which results in a large negative charge that is proportional to the RMM of the protein. The SDS-protein complex is then subjected to electrophoresis on a polyacrylamide gel. To aid visualisation the separated proteins are stained with silver or a dye such as coomassie blue to produce bands on the gel. By running standards of known RMM values on the

gel estimations of the RMM of unknown proteins can be made. The mass accuracy achievable by this method is typically no better than 10%.

Edman degradation is a traditional biochemical method of sequencing short peptides. Reaction between an isothiocynate compound and the N-terminal amine group of a peptide results in the sequential removal of the N-terminal amino acid of the peptide. The cleaved amino acid is normally characterised by HPLC. The shortened peptide is recycled to allow removal of the next amino acid. Practically Edman degradation is limited to approximately 30 consecutive residues as it is time consuming and requires relativity large amounts of sample.

Edman degradation also requires an unsubstituted amine at the N-terminus; the reaction is blocked by the presence of any substitutions. Furthermore, hydrophilic modifications such as phosphorylation, and glycosylation are not detected by this method.

1.2 (b) Higher order structure determinations

It has been suggested that the relative extent of charging observed in an ESI spectrum can be affected by the higher-order structure of a protein [46-50]. The high amount of charging typically observed for denatured proteins is believed to be due to the protein adopting a "loose" conformation thus exposing more of its potential charge accepting sites to the solvent. Under physiological pH conditions proteins form one or more unique conformations. Under such conditions some of the functional groups which may

normally be charged (i.e. N-terminus, arginine, lysine and histidine) may achieve electrostatic stabilisation by sharing a proton with another part of the protein. As a result, the net charge on proteins may depend on the conformation of the protein. It is considered that denatured proteins, i.e. in a open conformation, carry more charges than tightly bound proteins. If the charge state distribution detected by ESI-MS is related to the charge state of the protein ion in solution, then ESI-MS may be a potential probe of the solution conformation of proteins.

The high amount of charging typically detected for denatured proteins (usually electrosprayed in a solution of CH_3CN/H_2O in a 1:1 ratio plus 0.1 % HCOOH) can be altered by several factors; all deemed to be associated with the proteins' higher-order structure. For example the presence of disulfide bonds can lead to a lower amount of charging and thus a shift up field in m/z values; reducing the disufide bonds leads to an increase in the extent of charging [51,52]. Changing the pH of the solvent can also effect the extent of charging [47,49]. Analysing proteins in a 20 mM NH_4OAc (pH of approximately 7) solvent often results in lower charging. Changes in temperature and the addition of detergents have also been reported to affect the extent of charging [53].

Another method of probing protein conformational changes is to study the amide hydrogen exchange rates with deuterium. Conventional methods used to follow hydrogen exchange rates include nuclear magnetic resonance (NMR) [54] and ultra violet (UV) [55] spectroscopies. The combination of MS and hydrogen exchange illustrated that almost complete hydrogen exchange in bradykinin was detectable

[50,52]. Later studies demonstrated a large difference in the extent of exchange for a fully denatured protein and a folded protein, suggesting that a proportion of the hydrogen atoms in the latter were protected from exchange by the secondary and tertiary structure [52]. Other workers have combined MS and hydrogen exchange to study the non-covalent interaction between bovine α-lactalbuminn and the molecular chaperone GroEL [56].

ESI-MS can also provide information on tertiary structure by measurement of the collision cross section of an ion. Performed in a tandem quadrupole instrument, of quadrupole, hexapole, quadrupole (QhQ) geometry, these experiments involve introduction of ions of known translational energy into a collision cell of known gas pressure. As the ions traverse the cell they undergo multiple collisions, each of which reduces their translational energy. As the ions exit the collision cell their translational energies are measured, allowing calculation of the number of collisions they have undergone; this value is related to their collision cross section and hence conformation in the gas phase.

These studies indicate that cross sectional areas increase as the charge state of the ion increases, consistent with a more unfolded gas phase ion. These studies have been applied to small peptides (bradykinin and insulin) [57] and large proteins (serum albumin) [58].

Under standard ESI conditions multimeric proteins normally yield ESI spectra that

contain individual charge state distributions for each of the subunits present, as can be seen in chapter 5 for the soluble methane monooxygenase (sMMO) hydroxylase. Under certain conditions, however, various non-covalent interactions can be maintained in the gas phase and thus detected by means of mass spectrometry.

Several groups have shown that myoglobin, haemoglobin, and cytochrome c form complexes with protoporphyrin IX which can be detected by means of ESI-MS [47,48,59,67]. When myoglobin and haemoglobin were sprayed under acidic conditions no complex was observed; under aqueous conditions, at higher pH, complexes were detected. The cytochrome c complex is covalently bound and thus detected under either spraying condition. The pH of other spraying solvents has been found to be critical in the detection of other non-covalent complexes. Ganguly *et al* [61] demonstrated that pH and the presence of an organic modifier could be manipulated to produce peaks corresponding to a 21 kDa *ras* protein guanosine diphosphate (*ras*-GDP) complex, or only the *ras*-protein. At higher pH values (5-7) the complex is observed, whilst lower pH values resulted almost exclusively in the dissociation of the complex.

MALDI-MS has found fewer application in the study of non-covalent interactions. Suter *et al* [62] have demonstrated that non-covalent dimers of yeast alcohol dehydrogenase can be detected by MALDI. Because in a MALDI experiment the sample is dissolved in excess matrix, which may be denaturing, some non-covalent complexes cannot be observed.

Often MALDI-MS analysis of monomeric proteins yields peaks which correspond to multimers of the protein, thus care must be taken when assigning signals to non-covalent complexes.

1.3 References

1/ J. J. Thomson, *Ray of Positive Electricity*, Longmans, Green and Co., London (1913).

2/ A. O. Nier, *Rev. Sci. Instrum.*, **18**, 398 (1947).

3/ A. J. Dempster, *Phys. Rev.*, **18**, 415 (1921).

4/ M. S. B. Munson, F. H. Field, *J. Am. Chem. Soc.*, **88**, 2621 (1996).

5/ H. D. Beckey, *Int. J. Mass Spectrom. Ion. Phys.*, **2**, 500 (1969).

6/ E. W. Muller, *Phys. Rev.*, **102**, 618 (1956).

7/ M. Barber, R. S. Bordoli, R. D. Sedgwick, A. N. Tyler, *J. Chem. Soc. Chem. Commun.*, **7**, 325 (1981).

8/ W. Aberth, K. M. Strub, A. L. Burlingame, *Anal. Chem.*, **54**, 2029 (1982).

9/ R. M. Caprioli, T. Fan, J. S. Cottrel, *Anal. Chem.*, **58**, 2949 (1986).

10/ D. E. Games, S. Pleasance, E. D. Ramsy, M. A. McDowall, *Biomed. Environ. Mass Spectrom.*, **15**, 179 (1988).

11/ M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, *Int. J. Mass Spectrom. Ion Proc.*, **78**, 53 (1987).

12/ M. Karas, U. Bahr, F. Hillenkamp, *Int. J. Mass Spectrom. Ion. Proc.*, **92**, 231 (1989).

13/ F. Xiang, R. C. Beavis, *Rapid Commun. Mass. Spectrom.*, **8**, 199 (1994).

14/ J. Axelsson, A. M. Hoberg, C. Waterson, C. P. Myatt, G. L. Shields, J. Varney, D. M. Haddelton, P. J. Derrick, *Rapid Commun. Mass Spectrom.*, **11**, 209 (1997).

15/ C. K. Meng, C. N. McEwen, B. S. Larsen, *Rapid Commun. Mass Spectrom.*, **4**, 147 (1990).

16/ B. S. Larsen, C. N. McEwen, *J. Am. Soc. Mass Spectrom.*, **2**, 205 (1991).

17/ R. T. Gallagher, J. R. Chapman, M. Mann, *Rapid Commun. Mass Spectrom.*, **4**, 369 (1990).

18/ W. Paul, Z.Z. Steinwedel, *Naturforsch*, **8a**, 448 (1953).

19/ G. C. Stafford, P. E. Kelley, J. E. P. Syka, W. E. Reynolds, J. F. J. Todd, *Int. J. Mass Spectrom. Ion. Proc.* **60**, 85 (1984).

20/ J. C. Schwartz, M. E. Bier, D. M. Taylor, J. Zhou, J .E .P. Syka, M. S. James, G. C. Stafford, Proc. of the 43rd Conference on Mass Spectrometry and Allied Topics, Alanta, GA, May 21-26 (1995), pp.1114.

21/ E. M. Bier, J. C. Schwartz, J. Zhou, D. M. Talyor, J. E. P. Syka, M. S. James, B. Fies, G. C. Stafford, Proc. of the 43rd Conference on Mass Spectrometry and Allied Topics, Alanta, GA, May 21-26 (1995), pp.1117.

22/ J. N. Louris, R. G. Cooks, J. E. P. Syka, P. E. Kelley, G. C. Stafford, J. F. J. Todd,. *Anal. Chem.*, **59**, 1677 (1987).

23/ M. B. Comisarow, A. G. Marshall, *Chem. Phys. Lett.*, **25**, 282 (1974).

24/ K. D. Henry, J. P. Quinn, F. W. McLafferty, *J. Am. Chem. Soc.*, **113**, 5447 (1991).

25/ S. C. Beu, M. W. Senko, J. P. Quinn, F. M, Wampler, F. W. McLafferty, *J. Am. Soc. Mass Spectrom.*, **4**, 557 (1993).

26/ J. P. Speir, M. W. Senko, D. P. Little, J. A. Loo, F. W. McLafferty, *J. Mass Spectrom.*, **30**, 39 (1995).

27/ K. D. Henry, J. P. Quinn, F. W. McLafferty, *J. Am. Chem. Soc.*, **113**, 5447 (1991).

28/ L. Stryer, 'Biochemistry', 3rd Edition, W. H. Freeman and Company, New York, 1988, pp. 147.

29/ As for Ref. 28 pp. 16.

30/ As for Ref. 28 pp. 31.

31/ B. T. Chait, S. B. H. Kent. *Science*, **257**, 1885 (1992).

32/ A. Buzy, A. L. Millar, V. Legros, P. C. Wilkins, H. Dalton, K. R. Jennings, *Eur. J. Biochem.*, **254**, 602 (1998).

33/ K. Biemann, In 'Protein Sequencing: A Practical Approach', J. B. C. Findlay and M. J. Geisow (Ed.) Oxford University Press, 1989, pp.102.

34/ J. A Loo and R. R. Ogorzalek-Loo, In 'Electrospray Ionisation Mass Spectrometry', R. B. Cole (Ed), John Wiley and Sons, New York, 1997, pp. 395.

35/ K. Biemann, *Biomed. Environ. Mass Spectrom*, **16**, 99 (1988).

36/ T. P. Knepper, B. Arbogast, J. Schreurs. M. L. Deinzer, *Biochem.* **31**, 11651 (1992).

37/ A. M. Dizhoor, L. H. Ericsson, R. S. Johnson, S. Kumar, E. Olshevskya, S. Zozulya, T. A. Neubert, L. Stryer, J. B. Hurley, K. A. Walsh, *J. Biol. Chem.*, **267**, 16033 (1992).

38/ T. A. Neubert, R. S. Johnson, J. B. Hurley, K. A. Walsh, *J. Biol. Chem.*, **267**, 18274 (1992).

39/ S. A. Carr, M. J. Huddleston, R. S. Annan, *Anal. Biochem.*, **239**, 180 (1996).

40/ S. A. Carr, M. J. Huddleston, M. F. Bean, *Protein Sci.*, **2**, 183 (1993).

41. M. Wilm, G. Neubauser, M. Mann, *Anal. Chem.*, **68**, 527 (1996).

42/ K. Hirayama and S. Akashi, In 'Biological Mass Spectrometry: Present and Future', T. Matsuo, R. M. Caprioli, M. L. Gross and Y. Seyama (Eds.), John Wiley and Sons, Chichester, 1994, pp.299.

43/ P. A. Schindler, A. Van Dorsselaer, A. M. Falick, *Anal. Biochem.* **213**, 256 (1993).

44/ Y. P. See and G. Jackson in "Protein structure: A practical approach, T. E. Creigton, (Ed), IRL Press New York 1990.

45/ R. G. Krishna and F. Wold in "Protein structure: A practical approach, T. E. Creigton, (Ed) IRL Press New York 1990.

46/ J. A. Loo, R. R. Ogorzalek-Loo, H. R. Udseth, C. G. Edmonds, R. D. Smith, *Rapid Commun. Mass Spectrom.*, **5**, 101 (1991).

47/ Y. T. Li, Y. L. Hsieh, J. D. Henion, B. Ganew, *J. Am. Soc. Mass Spectrom.*, 4, 631 (1993)

48/ V. Katta, B. T. Chait, *J. Am. Chem. Soc.*, **113**, 8534 (1991).

49/ S. K. Chowdhury, V. Katta, B. T. Chait, *J. Am. Chem. Soc.*, **112**, 9012 (1990).

50/ V. Katta, B. T. Chait, *Rapid Commun. Mass Spectrom.*, **5**, 214 (1991).

51/ J. A. Loo, C. G. Edmonds, H. R. Udseth, R. D. Smith, *Anal. Chem.*, **62**, 693 (1990).

52/ V. Katta, B. T. Chait, *J. Am. Chem. Soc.*, **115**, 6317 (1993).

53/ J. C. Y. Le Blanc, D. Beuchemin, K. W. M. Siu, R. Guevremont, S. S. Berman, *Org. Mass Spectrom.*, **26**, 831 (1991).

54/ K. Wuthrich, NMR of proteins and nucleic acids, John Wiley and Sons, New York, 1986.

55/ J. J. Englander, D. B. Calhoun, S. W. Englander, *Anal. Chem.*, **92**, 517 (1979).

56/ C. V. Robinson, M. Grob, S. J. Eyles, J. J. Ewbank, M. Mayhew, F. Ulrich Hartl, C. M. Dobson, S. E. Radford, *Nature*, **372**, 646 (1994).

57/ A Method for the Determination of Collision Cross Sections of Gas Phase Ions by Means of Electrospray Ionisation Mass Spectrometry' A. C. Gill, K. R. Jennings, M. T. Bowers, presented at ASMS 1997.

58/ T. Covey, D. J. Douglas, *J. Am. Soc. Mass Spectrom.*, **4**, 616 (1993).

59/ R. Feng, Y. Konishi, *J. Am. Soc. Mass Spectrom.*, **4**, 638 (1993).

60/ M. Jaquinod, E. Leize, N. Potier, A. M. Albrecht, A. Shanzer, A. V. Dorsselaer, *Tet. Let.*, **34**, 2771 (1993)

61/ A. K. Ganguly, B. N. Dramanik, A. Tsarbopoulos, T. R. Covey, E. Huang, S. A. Furhman, *J. Am. Chem. Soc.*, **114**, 6559 (1992).

62/ M. J. F. Suter, W. T. Moore, T. B. Farmer, J. S. Cottrell, R. M. Caprioli, In 'Techniques in protein chemistry III', R. H. Angeletti, (Ed),Academic press 1992.

CHAPTER 2

INSTRUMENTATION, IONISATION AND TANDEM MASS SPECTROMETRY

2.1 Introduction

The vast majority of the electrospray ionisation m/z spectra presented in this work were acquired using a Micromass Quattro II tandem mass spectrometer of quadrupole, hexapole, quadrupole configuration. Where greater sensitivity and resolving power were required a Micromass Q-ToF instrument was used. The Q-ToF mass spectrometer is a hybrid instrument of quadrupole, hexapole, orthogonal acceleration time of time mass analyser configuration [1].

This chapter will discuss the formation of gas phase ions by electrospray ionisation as well as the methods employed to separate ions according to their mass/charge ratios. Ion detection and low energy tandem mass spectrometry are also discussed.

2.2 The formation of gas phase ions by means of electrospray ionisation mass spectrometry

ESI can be described as the electrostatic nebulisation of a charged analyte solution, followed by the evaporation of the resultant droplets, to produce singly and multiply charged gas phase ions of the type $[M+nH]^{n+}$ from a sample having a RMM of M. ESI promotes little fragmentation of the ions it produces and in some situations has been shown to retain the solution phase conformation of the analyte in the gas phase [2,3]. For these reasons ESI is termed a "soft" ionisation technique.

The seminal research into the use of ES as an ionisation method was performed by Dole et al [4,5] in 1968. Dole's experiments consisted of pumping a dilute polystyrene solution through a hypodermic needle into a cylindrical chamber through which flowed a nitrogen bath gas at atmospheric pressure. A potential of several kilovolts was maintained between the needle and the chamber wall. The resultant electric field at the needle tip dispersed the liquid being pumped through it into a fine spray of charged droplets.

Dole believed that the electrosprayed analyte solution would result in droplets containing only one analyte ion. The constraint, however, of not having available a mass analyser which could accommodate ions of the molecular weight range Dole was investigating meant that direct determinations of m/z were not feasible.

Dole, therefore, calculated the kinetic energy of the ions by measuring the retarding potential (V_r) required to keep certain light ions from reaching a Faraday cup collector. The ratio of m/z values were then calculated, assuming that during the free jet expansion into vacuum the ions were accelerated to the readily calculable terminal velocity of the bath gas, using the equation 2.1.

Using this method Dole also presented experimental evidence for the ionisation of a protein (Zein, 50,000 Da) [6]. Dole believed that ions were formed over a series of stages until finally, one ion would occupy one solvent droplet. This idea formed the basis of the 'charge residue' or 'single ion in a droplet' model.

$$\frac{mv^2}{2} = zeV_r \qquad \text{Equation 2.1.}$$

where, m = Mass of ion.

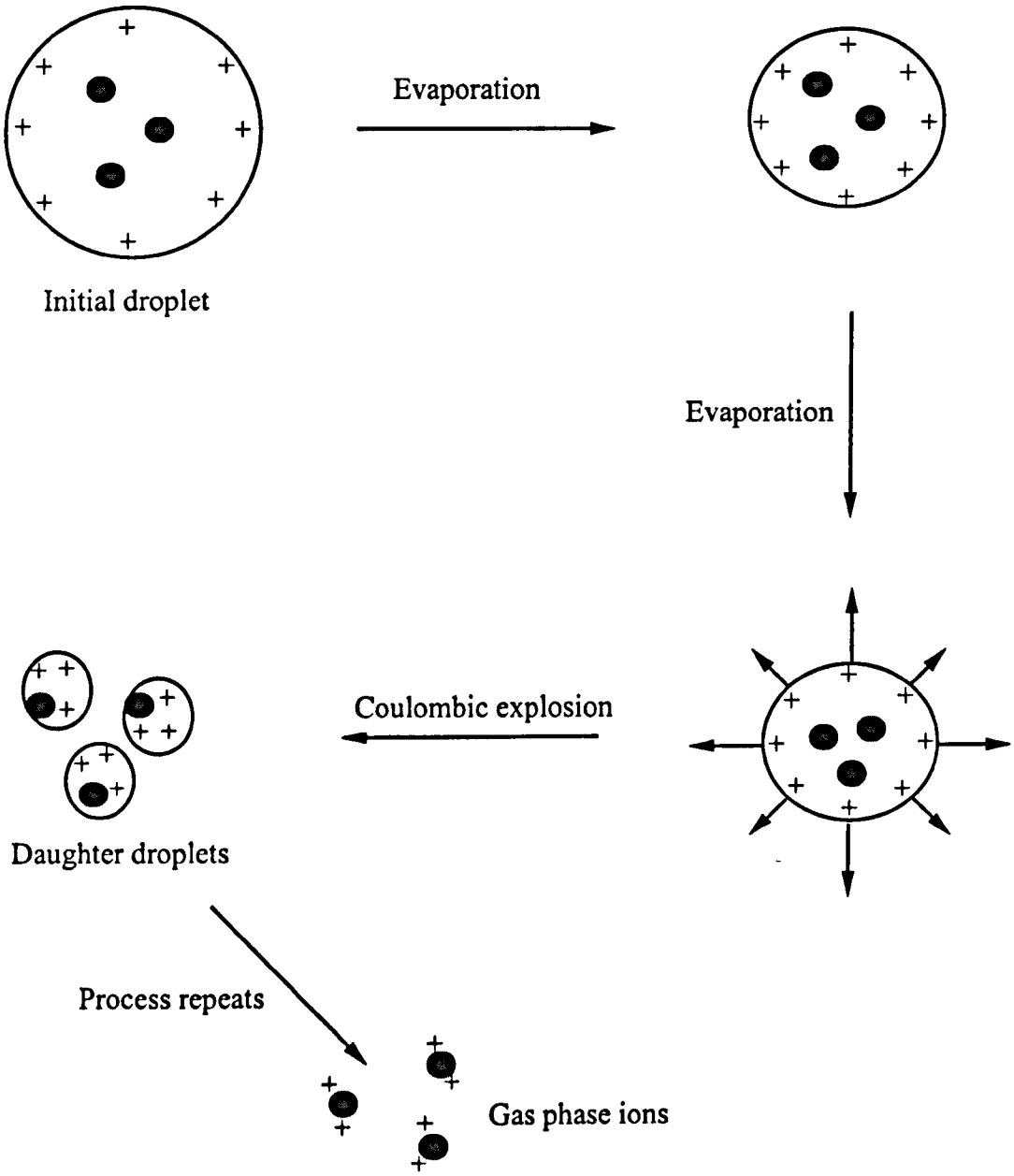
v = Velocity.

z = Number of charges.

Vr = Retarding voltage.

e = Electronic charge.

This model suggests that the highly charged droplets are initially stable as the surface tension of the droplet is greater than the coulombic repulsion between the charges. As the droplet is subjected to evaporation the charge density on the surface increases until the coulombic repulsion between ions on the droplet surface becomes equal to the surface tension. At this point the Rayleigh stability [7] limit has been reached and further evaporation leads to the violent disruption of the droplet in a process that is often referred to as a coulombic explosion. This leads to the production of smaller daughter droplets which undergo an identical process. Continued evaporation and subsequent explosion of the droplets is thought to result in the dispersion of gas phase ions in the bath gas. This process is illustrated in scheme 2.1. A portion of this gaseous dispersion could then be expanded as a supersonic free jet into the reduced pressure of a mass analyser.



Scheme 2.1. Schematic representation of the "Charge residue" or "single ion in a droplet" model as proposed by Dole. The schematic illustrates the evaporation and subsequent explosion of the initial droplet. Further evaporation and explosions are suggested to lead to the production of gas-phase ions.

More recently, an alternative mechanism, field desorption has been proposed by Iribarne and Thomson [8]. They suggested that as the evaporating droplet radius decreases the electrostatic field at the surface increases. An analyte molecule that has accumulated sufficient charge is then able to desorb from the charged droplet alone or in association with one or more molecules of solvent as illustrated in scheme 2.2. In subsequent studies Thomson and Iribarne presented mass spectrometric analysis of the desorbed ions [9].

The thermodynamics of the field desorption process have been investigated by Sakairi and Yergey [10] and used successfully to predict trends in molecular ion intensities for amino acids analysed by ESI-MS. Rollgen *et al* [11] however have questioned the Thomson and Iribarne theory, as they believe that field desorption cannot occur before the Rayleigh stability limit has been reached. They reason that a Rayleigh explosion would occur before field desorption could take place. No one mechanism has yet been accepted universally as the method of production of gas phase ions from solvent droplets.

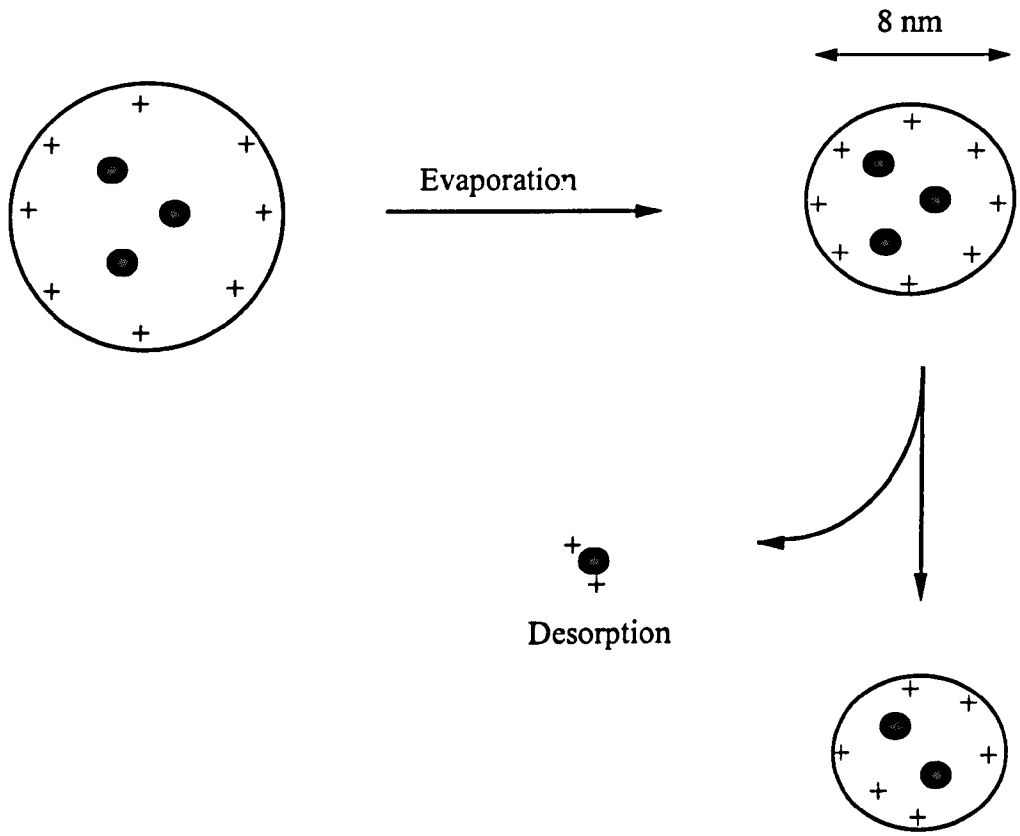
The development of ESI as a mass spectrometric technique occurred in 1984. Yamashita and Fenn [12] demonstrated ESI-MS in the negative ion mode using a quadrupole mass analyser of limited m/z range and later also investigated the interface of HPLC and ESI-MS [13].

One of the fundamental features of electrospray ionisation is its ability to produce multiply charged ions. It is likely that Dole observed multiply charged ions in some of his early experiments, but misinterpreted the charge states. In 1985, Fenn and co-

workers [13] observed dominant contributions for the doubly charged polypeptide gramicidin S (1141 Da). More significant was the observation by Fenn and co-workers of poly(ethylene glycol) samples of average molecular weights of up to 17,500 Da with up to 23 charges present [14].

The broad molecular weight distributions of the samples, however, meant that only unresolved peaks were visible in the m/z range of 500-1400 in the mass spectra. Subsequently, the methodology was applied to large biopolymers, which overcame this problem by having well defined molecular weights.

Fenn [15] observed multiply protonated molecular ions of up to 40,000 Da having as many as 45 charges. With this development came the realisation that ESI-MS could be applied to a diverse range of biological problems.



Scheme 2.2. Schematic representation of the "Ion evaporation" model as proposed by Thompson and Irvine.

2.3 The Micromass Quattro II tandem mass spectrometer

The Micromass Quattro II is a tandem mass spectrometer of quadrupole hexapole quadrupole (QhQ) configuration. As a tandem instrument it can be used for both MS and MS/MS analysis. For MS analysis ions are mass analysed in the first quadrupole

(Q1) and then diverted off axis by a high voltage lens ('Altrincham' lens) to a photomultipler for detection as shown in Figure 2.1.

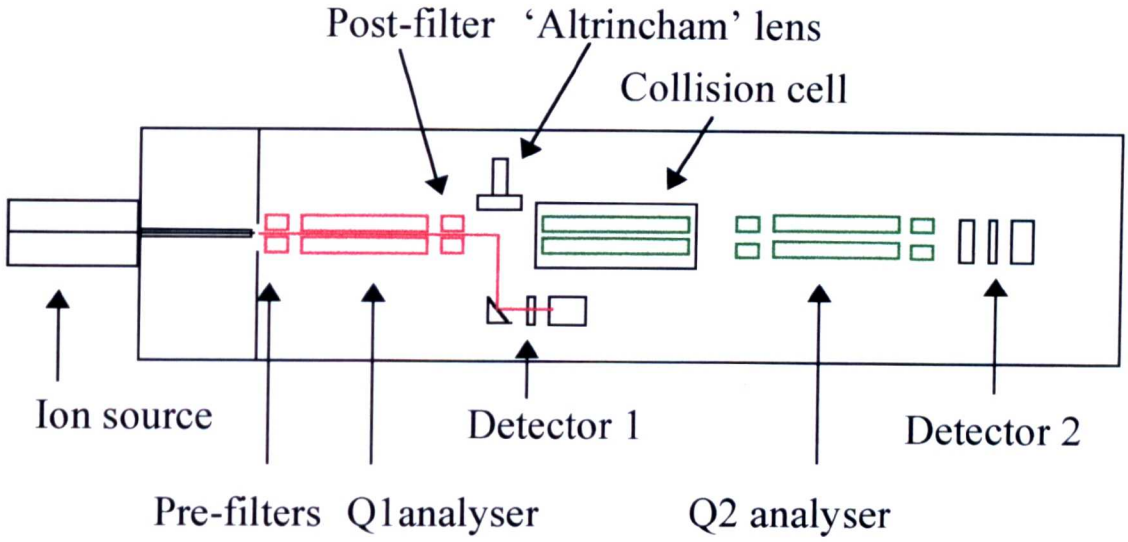


Figure 2.1. Schematic representation of the configuration of the Quattro II tandem mass spectrometer operating in MS mode.

For MS/MS analysis parent ions are mass selected by Q1 and then accelerated into the RF only gas filled hexapole collision cell. After undergoing CID the fragment ions are passed into the second quadrupole (Q2) for mass analysis and detection by a second photomultipler situated after Q2. The Quattro II is available with a wide variety of ionisation sources including EI, CI, electrospray, APcI and thermospray. Electrospray ionisation was used throughout this work.

The standard flow electrospray ionisation source of the Quattro II instrument consists of a nebuliser, heated high voltage (H.V.) lens, ion sampling cone and skimmer assembly. A schematic of the electrospray source is shown in Figure 2.2. The sample is infused by a syringe driver and is transported to the source through a fused silica capillary.

The solvent is typically composed of a volatile organic substance (acetonitrile or methanol) and water in a 1:1 mix.

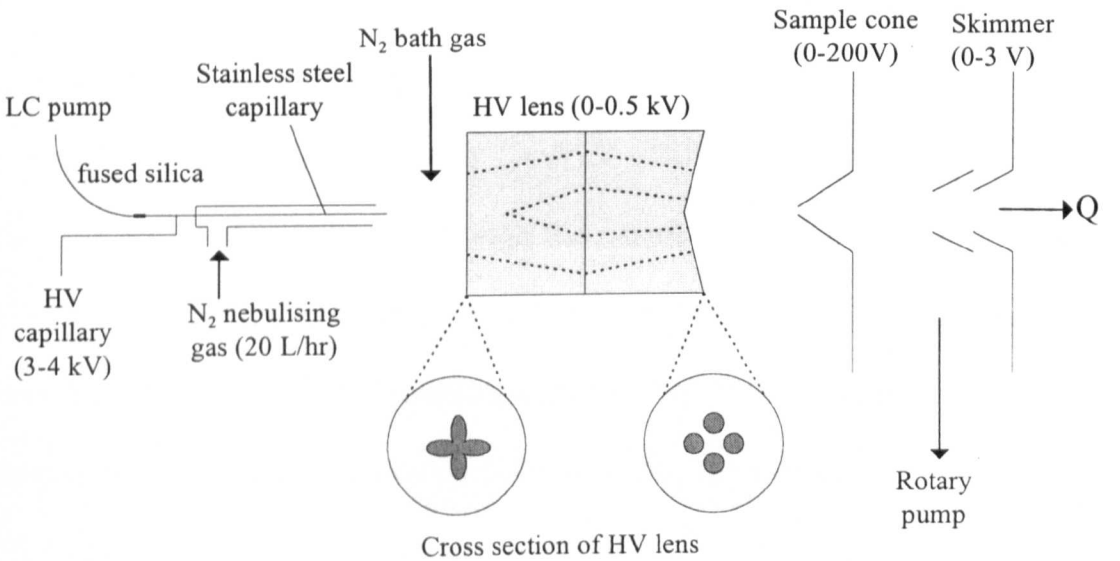


Figure 2.2. A schematic of the "Quattro II" electrospray source.

To aid protonation, in some applications, the solvent is acidified by the addition of 0.1 - 1% of an acid e.g. formic acid.

Depending upon solution properties and the field configuration the sample solution emerges from the source needle as a Taylor cone [16]. At higher capillary potentials the cone becomes unstable and a liquid filament with a diameter of a few micrometers is emitted from the cone tip. At some point the filament becomes unstable and breaks up to form droplets. The formation of the droplets can occur in one of several different modes [17].

The H.V. lens is shaped such that it captures a large angle of the electrospray plume and focuses the ions towards the sampling cone. A further important point is that the H.V. lens reduces the amount of involatile material or liquid from entering the mass analyser, thus minimising any possible contamination of the instrument.

The ions and droplets then pass through the sample cone, which contains a small (0.2 mm) aperture leading into the reduced pressure region of the mass analyser. The reduction from atmospheric pressure to approximately 10^{-1} mbar (provided by rotary pumps) results in a free jet expansion of the ions and gas molecules.

Finally the skimmer assembly minimises the number of any solvent molecules entering the mass analyser. Further pumping after the skimmer assembly (provided by turbo-molecular pumps) reduces the pressure to a value of approximately 10^{-5} mbar.

As mentioned earlier, an important characteristic of the electrospray ionisation process is that the ions produced are often multiply charged. This is an advantageous situation, as the effective mass range of the mass analyser used is increased considerably. Very large compounds (greater than 100,000 Da) can therefore be detected by a mass analyser of relatively modest m/z range when using an ESI source. For this reason the quadrupole mass analyser is commonly used with ESI. The quadrupole also has the advantage of being simply interfaced to the ESI source and being relatively inexpensive.

2.3 (a) The quadrupole mass filter

The Quattro II instrument utilises two high performance, research grade quadrupole mass analysers, each incorporating a pre-filter assembly in order to reduce fringe field effects and to protect the main analyser from contaminating deposits.
The quadrupole mass filter consists of four rods which are parallel to the *z*-axis and perpendicular to the *x* and *y*-axis. Theoretically each rod should have a hyperbolic cross-section but, in practice, cylindrical rods are satisfactory if properly spaced.

Two opposed rods have a potential of $+(U + V \cos \omega t)$ and the other two, $-(U + V \cos \omega t)$ where U is a fixed potential and $V \cos (\omega\tau)$ represents a radio frequency (RF) field of amplitude (V) and frequency (ω) as illustrated in Figure 2.3. The stability and paths of ions within a quadrupole mass filter, in which the separation of the opposite rods is $2r_o$, can be calculated from the Mathieu equations [18].

$$a = \frac{8eU}{mr_0^2\omega^2}$$

Equation 2.2.

$$q = \frac{4eV_0}{mr_0^2\omega^2}$$

Equation 2.3.

hence,

$$\frac{a}{q} = \frac{2U}{V_0}$$

Equation 2.4.

In the transverse direction of the quadrupoles an ion will oscillate amongst the poles in a complex fashion depending on its m/z value and the values of a and q. By suitable choice of a and q values it can be arranged that only ions of a particular m/z ratio will oscillate in a stable manner about the central axis; in this case all other ions will oscillate with greater and greater amplitude until eventually they strike the rods or pass between the rods and are lost.

The description of path stability is represented schematically in the form of a stability diagram as shown in Figure 2.4.

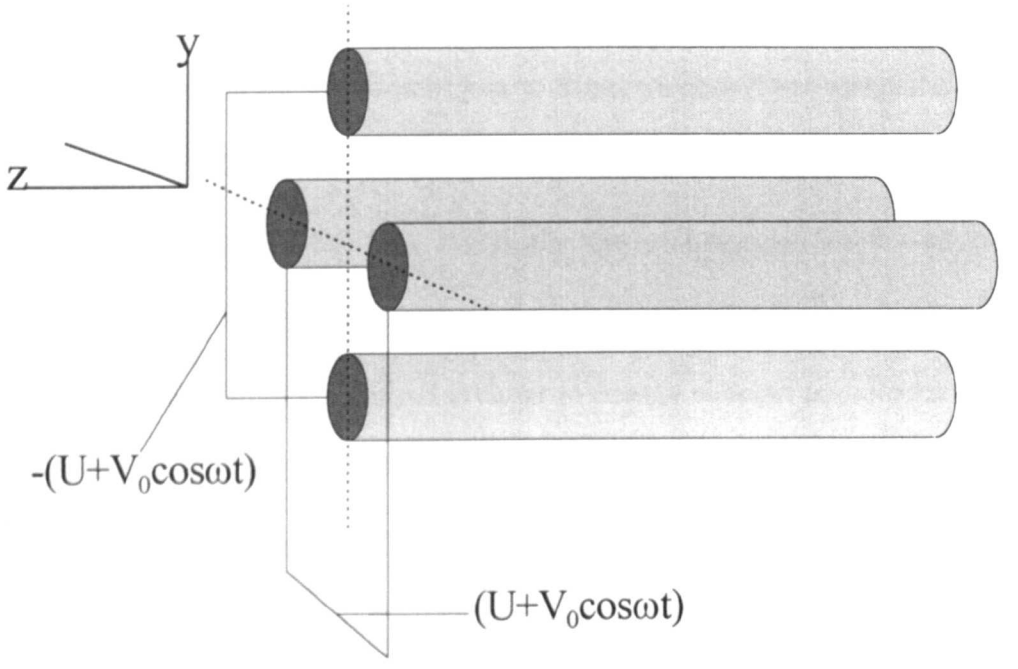


Figure 2.3. Quadrupole mass filter rod assembly showing the application of +(U + V cos ω t) to one pair of opposing rods and the -(U + V cos ω t) to the adjacent pair of rods.

In Figure 2.4. the ion represented by m_1 lies within the region of stable oscillations and will therefore pass through the mass filter. The ion represented by m_2 lies outside of the area of stable oscillations for these particular values of a and q and will not pass through the mass filter.

Where the straight line (2) of slope a/q enters and leaves the region of stability defines an ion of a particular m/z which will be transmitted through the mass filter. For line 3 of smaller slope (smaller a/q) a larger range of m/z values will have stable paths through the mass filter resulting in decreased resolution and increased sensitivity [19]. Conversely, for a line (1) which passes through the apex (R) of the

region of stability, no ions of any m/z are transmitted. To ensure that only ions of any one selected m/z value are transmitted (i.e. maximum resolution), the ratio a/q must be chosen such that the line passes close to R but which still lies within the region of stability. This will give maximum resolution for the instrument, although this will be at the cost of reduced sensitivity. Practically the resolution attainable with a quadrupole mass filter is limited due to the low ion velocities and the positions at which ions enter the filter [20]. It is easier to change voltages than frequencies, therefore to transmit ions of other m/z values the frequency is kept constant but the voltages are varied in such a way that a/q remains constant. By continuously increasing or decreasing a and q whilst keeping a/q constant, ions of increasing or decreasing m/z successively traverse the quadrupole assembly to give a m/z spectrum.

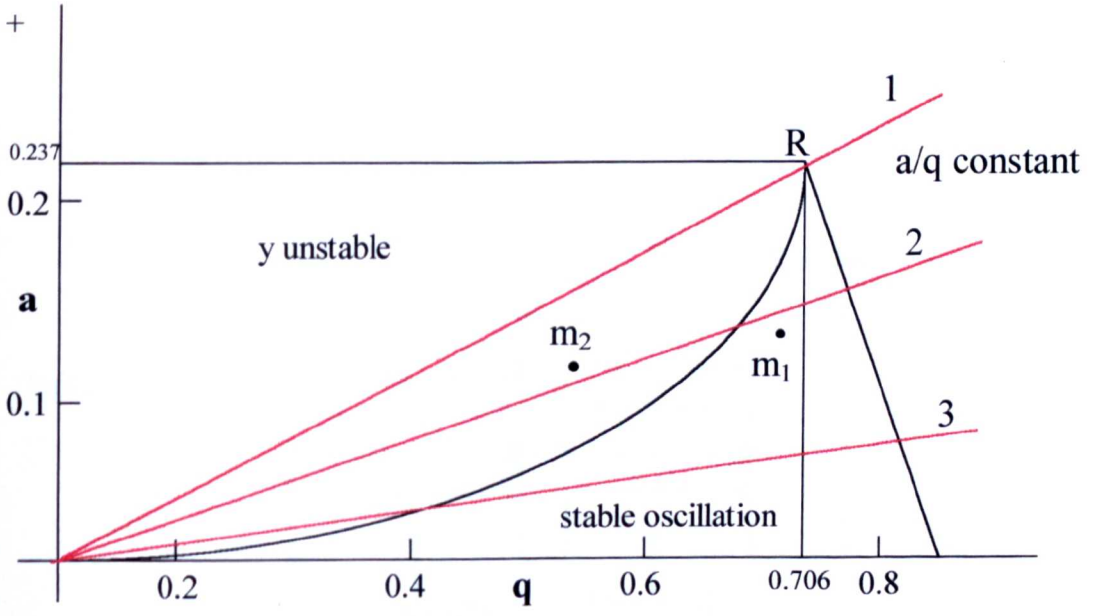


Figure 2.4. Stability diagram showing mass separation according to various values of a and q.

2.3 (b) The Quattro II Ion detection assembly

The Quattro II uses two photomultipliers, one situated off axis after Q1 and the other on axis after Q2. The principle of ion detection by the conversion of an ion current first to electrons, then a flux of photons then to a current of electrons which is amplified by a photomultiplier [21], is shown in Figure 2.5. The arrival of a fast ion upon a dynode causes the emission of electrons which are accelerated onto a phosphor screen operated at approximately 10 kV. This incidence results in the generation of photons. The emitted light is detected by a photomultiplier and is converted into an electric current. In the Quattro II the photomultiplier is sealed in a vacuum tube to protect it from analyte or pump oil vapours.

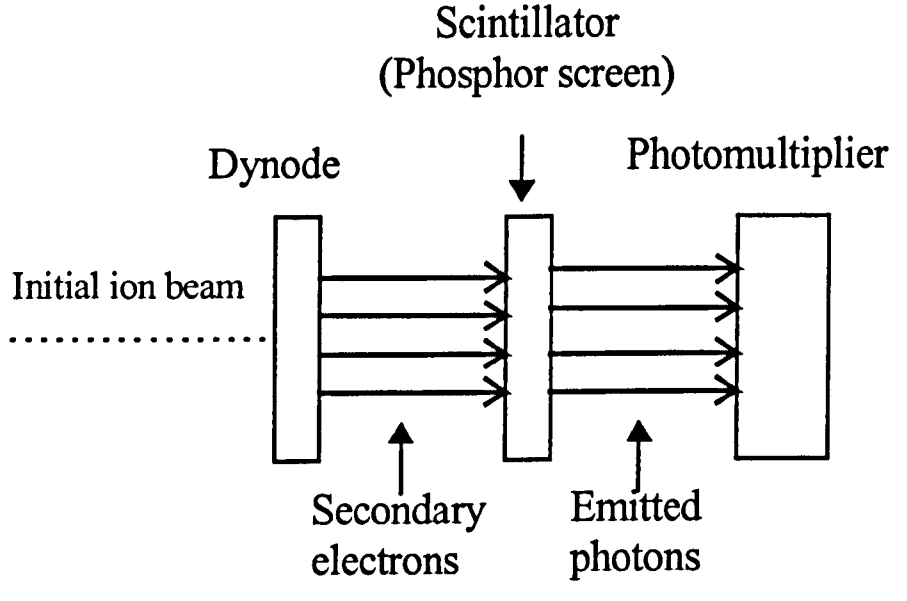


Figure 2.5. Schematic of a photomultiplier detector

2.3 (c) MassLynx data system

The Quattro II instrument is operated via by the windows based MassLynx NT data system. The software allows all source and analyser voltages to be adjusted on screen. The software is also used for data processing and data base searches. One particularly useful feature of MassLynx is the maximum entropy ('MaxEnt') deconvolution algorithm [22]. The 'MaxEnt' program deconvolutes multiply charged ESI data to a true mass scale. This is especially helpful when analysing mixtures of proteins of peptides or confirming the presence of minor components.

2.4. The Micromass Q-ToF quadrupole orthogonal acceleration time of flight mass spectrometer.

The Micromass Q-ToF instrument, like the Quattro II, is a tandem mass spectrometer capable of both MS and MS/MS analysis. It differs from the Quattro II in that it employs an orthogonal acceleration time of flight mass analyser, which is used in both MS and MS/MS modes of operation of the type first described by Dodonov *et al* [23]. For MS analysis ions are passed through a RF quadrupole mass filter to a RF hexapole collision cell, which in MS mode is not filled with collision gas, to a two stage orthogonal acceleration single stage reflectron time of flight mass analyser as shown in Figure 2.6.

When operating in MS/MS mode the quadrupole and hexapole act in an identical manner to those in the Quattro II instrument i.e. the quadrupole selects the parent ion whilst the hexapole acts as a collision cell. The benefits of using the ESI orthogonal TOF configuration is that continuous and simultaneous detection of ions across the

100 fold gain in sensitivity over conventional scanning mass spectrometers. The instrument also has a m/z range in excess of 10,000 [24].

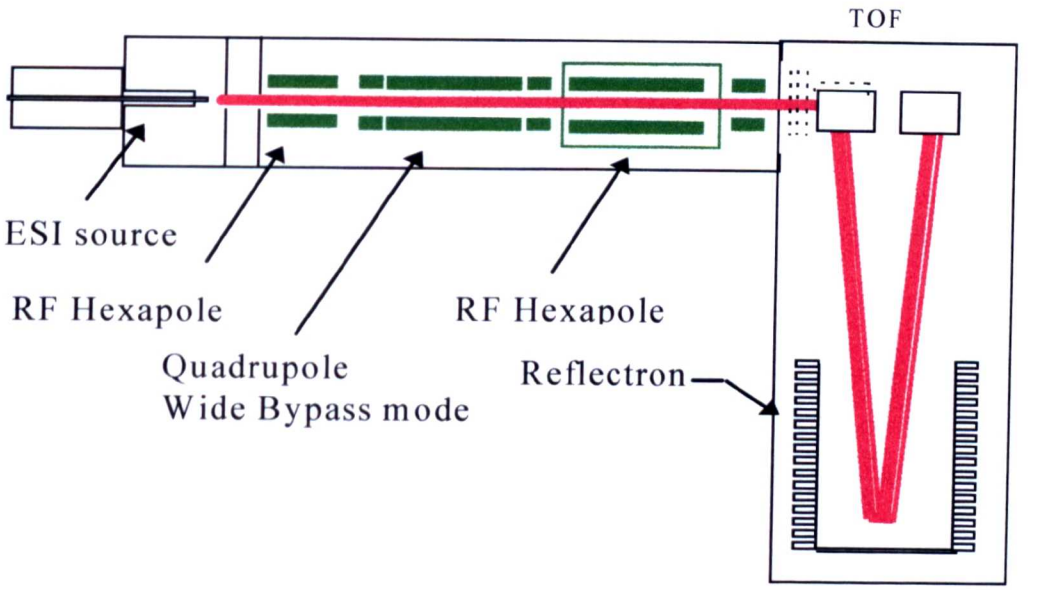


Figure 2.6. Configuration of the Q-ToF instrument operating in MS mode.

2.4 (a) Time of flight mass analyser

In a time of flight mass spectrometer, ions formed in an ion source are extracted and accelerated to a high velocity by a very short pulse of ~25 kV into an analyser consisting of a long straight drift tube. The ions travel along the drift tube until they reach a detector. After the initial acceleration phase the velocities reached by the ions are inversely proportional to the square root of their m/z values.

It is necessary to extract small pulses of ions because, using only time to differentiate amongst the masses, it is important that the ions all leave the ion source at the same instant with the same kinetic energy. The first step is acceleration through an electric field (E volts). The kinetic energy of the ion is given by equation 1.

$$\frac{mv^2}{2} = z.e.E$$

Equation 2.5.

Equation 2.5.follows by rearrangement,

$$v = \sqrt{(2z.e.E / m)}$$

Equation 2.6.

If the distance from the ion source to the detector is d, then the time (t) taken for an ion to traverse the drift tube is given by the equation

$$t = d / v = d / \sqrt{(2e.z.E / m)} = d.[\sqrt{(m / z)}] / \sqrt{(2e.E)}$$

Equation 2.7.

In equation 3 d is fixed, E is held constant in the instrument and e is a universal constant. Thus, the flight time of an ion t is directly proportional to the square root of m/z (equation 2.7).

$$t = \sqrt{(m / z)} \times \text{ a constant}$$

Equation 2.8.

The resolution attainable on a typical TOF instrument is in general limited, particularly so at higher mass. There are two major reasons for this; first the flight times are inversely proportional to the square root of m/z thus as the mass/charge ratio increases the difference in arrival times between ions of similar m/z values decreases and so become increasingly more difficult to differentiate. The second reason is that not all ions of any given m/z value will reach exactly the same velocity or they are not all formed at the same point in the ion source. Thus even for one m/z

value the ions reach the detector over a short interval of time rather than all at once. When the separation times of flight times is very short the spread for individual m/z values means that there will be an overlap in arrival times between ions of closely related m/z values.

Two methods exist for improving the resolution of a time of flight analyser. The first is based on the work of Wiley and McLaren who developed the time-lag focusing technique [25]. Using this method the ions are allowed a period of delay before the accelerating voltage is applied to them. This has the effect of allowing ions of identical mass but different kinetic energies to move away from each other. Thus, the ions which move further away from the extraction plates experience a greater potential than those which initially move towards the extraction plates. Consequently, the former ions have a greater kinetic energy and hence a greater velocity than the latter ions in the drift tube. This results in the focusing of ions with identical m/z values at the same point thus reducing the kinetic energy spread and improving resolution.

Another method for improving the resolution of a time of flight analyser is to make use of a reflectron [26]. A reflectron is a homogenous electrostatic field placed at the end of the flight path of the ions and has a polarity the same of that of the ions i.e. positive ions experience a retarding positive potential. Upon entering the reflectron the ions come to a stop and are then accelerated in the opposite direction.

The reflectron is at a slight angle to the line of the flight of the ions, and when reflected, the ions do not travel back along the same path but along a slightly deflected one. For ions of any given m/z value the ones with greater kinetic energy and hence velocity will have shorter drift times but will travel further into the electrostatic field than the slower ions before being reflected. As a result the faster ions spend more time than the slower ones within the reflectron. This focuses the faster and slower ions so that although they leave the reflectron still with different velocities they arrive at the detector together and the resolution is greatly improved. There are disadvantages to using the reflectron in that the sensitivity of the instrument is decreased through ion loss by collision and the second length of analyser. These problems are more severe for ions of large mass and for these reasons the reflectron is often not used in these circumstances.

TOF MS is a natural choice for pulse ionisation sources, the most important being MALDI as it provides a complete spectrum for each event. There are however difficulties in coupling a continuous ionisation source such as ESI to a TOF mass analyser along the spectrometer axis without a large loss in sensitivity [27]. Fortunately a TOF analyser can tolerate a relativity large spatial or velocity spread in a plane perpendicular to the spectrometer axis. The ions produced by an ESI source can thus be injected into the TOF analyser perpendicular to the axis, i.e. orthogonal injection [23, 28-30]. This geometry provides a high efficiency interface for transferring ions from a continuous beam to a pulsed method. A second advantage is the small velocity spread in the z direction makes high resolution easier to obtain.

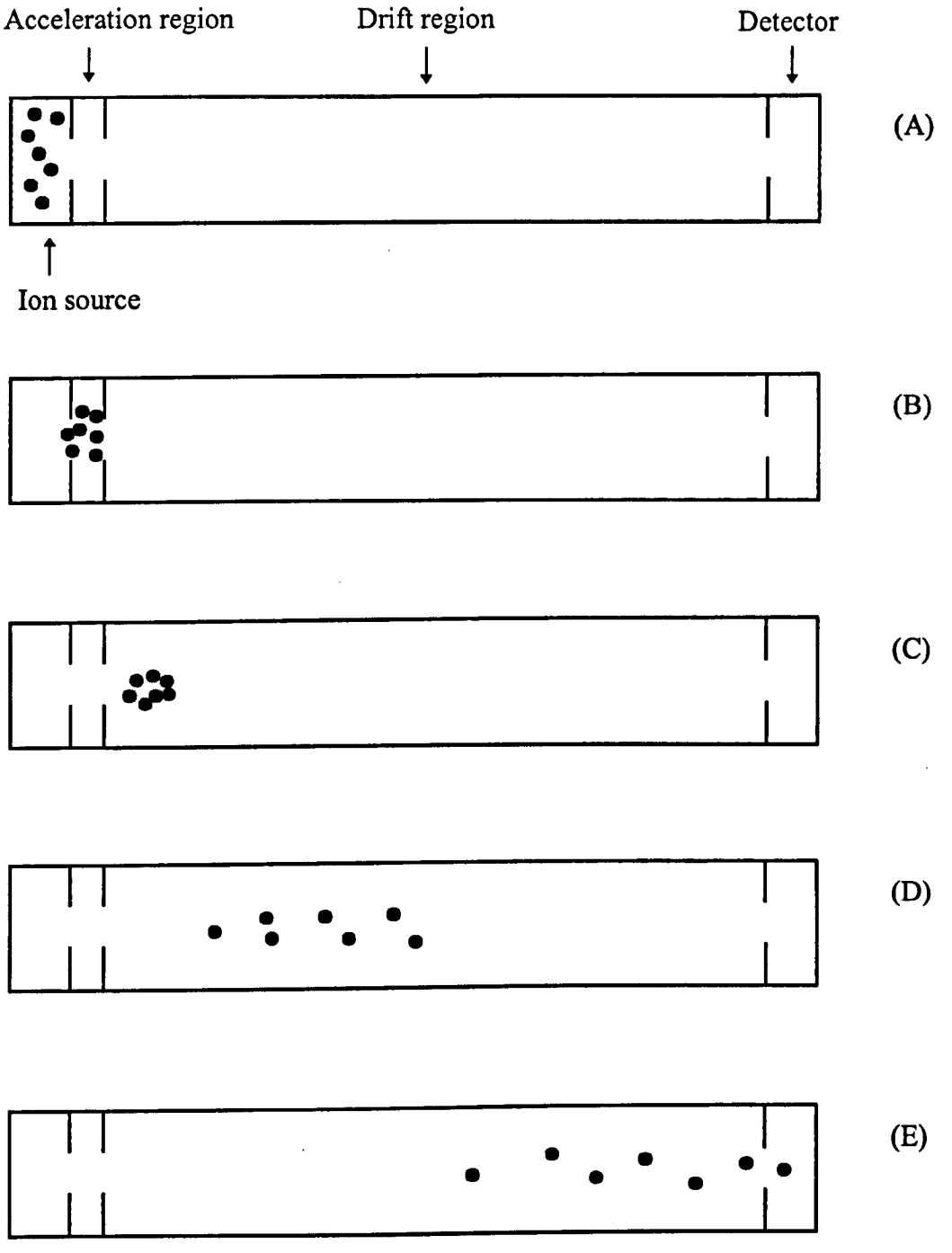


Figure 2.7. Time of flight optics in a linear instrument. A pulse of ions are formed in the ion source (A) and then accelerated out of the source (B) into the drift region (C). After a short while the ions have separated along the drift region according to their m/z value. In (E) the ions with the smallest m/z value arrive at the detector, followed by ions with increasing larger m/z values.

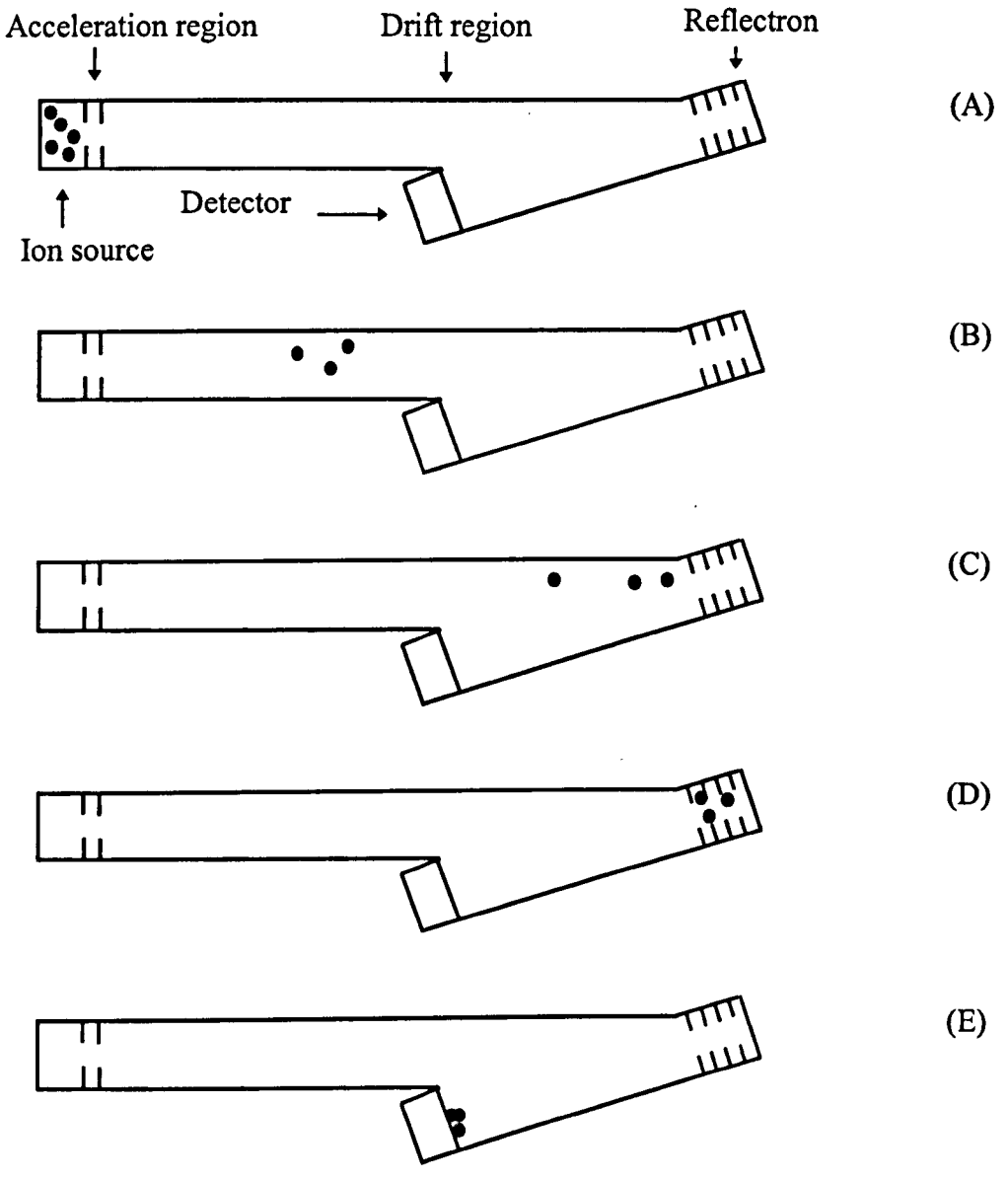


Figure 2.8. Time of flight optics in an instrument equipped with a reflectron. Ions are formed in the ion source (A) of the same m/z value. In (B) the ions have been accelerated but, some having different kinetic energies, they have spread out in space. In (C) the ions approach the reflectron, which they penetrate to different depths depending on their kinetics energies(D). In (E) the focused ions, still with different velocities, reach the detector close together.

2.4 (b) Ion detection in the Q-Tof instrument

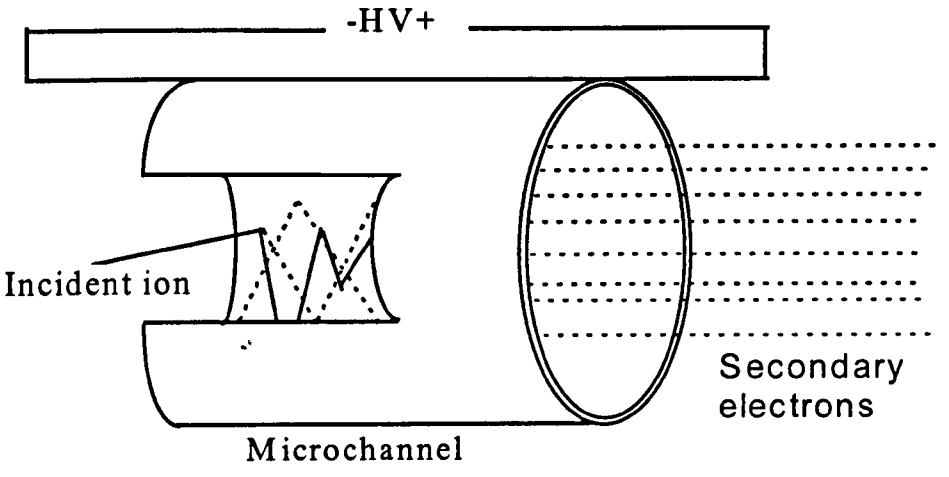


Figure 2.9. Production of secondary electrons in a MCP.

The Q-ToF instrument employs a dual stage microchannel plate detector (MCP). The MCP consists of an array of lead glass tubes in a solid assembly. The assembly is sliced into wafers at an angle and polished to an optical finish. The solid cores are removed to leave porous structures of many thousands of uniform tiny holes or microchannels. Each microchannel acts as a channel electron multiplier. When an ion strikes the input of the channel wall, secondary electrons are generated and accelerated down the channel wall and additional secondary electrons are produced. A high voltage of up to 1kV is applied across the microchannel plate, the current through the semiconductive layer supplies the electrons necessary to produce the secondary electrons. Due to the high number of microchannels used in a MCP very high gains are attained [31]. Further advantages of the MCP are that they are compact, stable when exposed to air and have good signal to noise characteristics.

2.5 Tandem mass spectrometry

As mentioned at the beginning of the chapter ESI is a relatively soft ionisation technique as it promotes very little fragmentation and this is an advantageous situation when RMM information is required. It provides, however, very little or no structural or sequence information. This type of information about an ion can be obtained by subjecting it to a tandem mass spectrometry (MS/MS) experiment. These experiments involve the selection of the ion of interest, often called the parent ion, in a first mass analyser. This ion is then subjected to CID to produce structurally informative fragment or daughter ions which are separated according to their m/z values in a second mass analyser and subsequently detected. This methodology is a valuable way of characterising unknown compounds or confirming the presence of a known compound. MS/MS analysis can be especially useful when selecting a particular ion from a complex mixture as this adds a high degree of specificity.

ESI-MS/MS has been demonstrated to be a suitable method for sequencing peptide ions [32] especially tryptic peptide ions [33,34] which are often doubly charged, thus fragmenting to produce singly charged ions which are relatively easy to assign. Fragmentation can also be produced in the ESI source by increasing the voltage between the atmosphere-vacuum aperture and the first skimmer (called the cone voltage on the Quattro II instrument). In this region, the pressure is such that multiple collision between ions and neutrals will occur. The applied voltage increases the energy imparted to the ions during the collisions with neutral molecules which can result in fragmentation This type of fragmentation provides no mass selection and as such works best with reasonably pure compounds, although it does

provide the facility to perform MS/MS/MS (MS^3) on standard tandem instruments. Workers have successfully applied this method of fragmentation to peptide sequencing and the identification of small molecules [35,36,37].

In the Quattro II instrument used in this study Q1 is used to select the parent ion which is passed into the hexapole collision cell containing argon gas at a typical pressure of approximately 5×10^{-3} mbar. The collision cell is a RF only device (i.e. it does not act as a mass filter) which, once the precursor ion has undergone CID, acts to refocus the ions scattered by the multiple collisions. A positive ion trajectory within an RF only hexapole will be towards the rod with a negative potential applied to it, as the potential rapidly changes to positive the ion will be repelled back to a central position within the rods [38]. In this manner fragment ions are focused for maximum transmission.

A DC potential (the collision voltage) is applied to all six rods of the hexapole relative to Q1. The polarity is such that the precursor ion is accelerated into the collision cell. When operating in MS/MS mode all electrodes after the collision cell automatically have the collision voltage algebraically added to their set voltages. This is to compensate for the fact that the ions at the exit of the cell have essentially zero energy as a consequence of the multiple collisions they have undergone and thus have to be extracted.

The configuration of the Quattro II when operating in MS/MS analysis is shown in Figure 2.10.

MS/MS analysis using the Q-ToF instrument is performed in a manner similar to that used with the Quattro II instrument. The first quadrupole acts to select the precursor ion which is then accelerated into the collision cell where it undergoes CID upon impact with the collision gas. The fragment ions are mass analysed by the TOF analyser as shown in Figure 2.11.

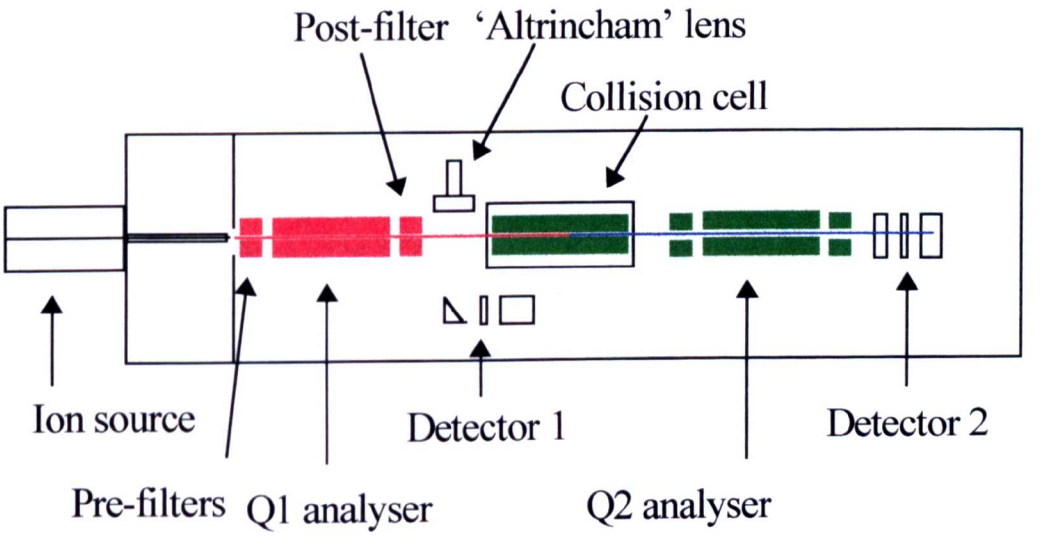


Figure 2.10. Configuration of the Quattro II instrument for MS/MS analysis.

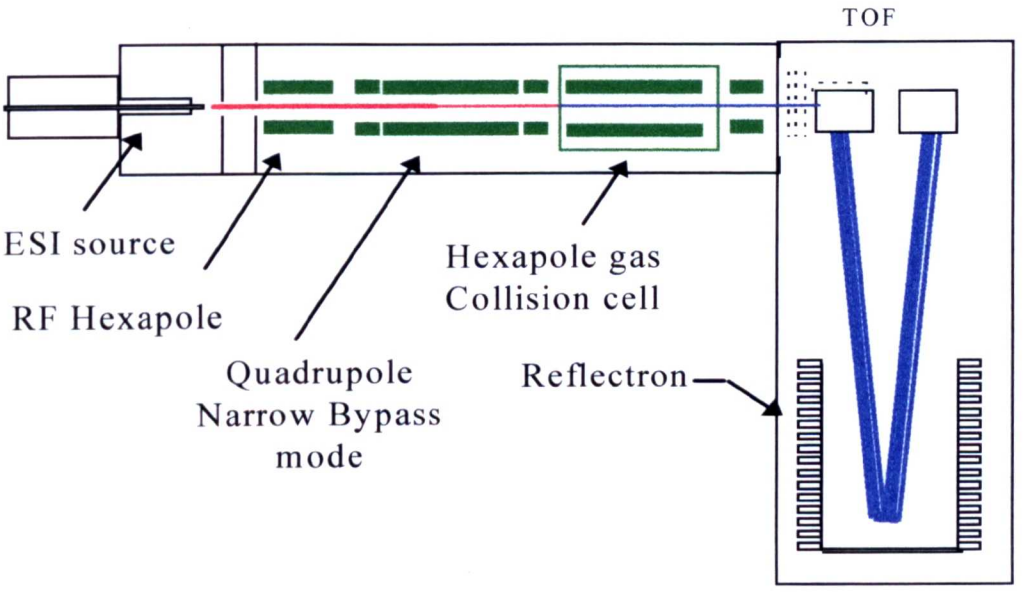


Figure 2.11. Configuration of the Q-ToF instrument when operating in MS/MS mode.

2.5(b) Low energy collision induced decomposition

Since being first described in 1968 [39] CID has become an accepted and routine analytical method. The majority of low energy CID is performed in tandem quadrupole instruments where the collision cell itself is typically a hexapole or a octopole. The first tandem quadrupole instrument designed for analytical applications was described by Yost and Enke [40,41]. Typically modern tandem quadrupole instruments offer collision energies up to approximately 200eV for singly charged ions and an m/z range of approximately 4000.

The process of collision induced decomposition (CID) of an ion to produce structurally informative fragment ions is believed to occur via a two step process [42,43]; activation of the ion which occurs during the interaction between the accelerated precursor ion and the target gas, followed by unimolecular dissociation. As the accelerated precursor ion strikes the stationary collision gas a proportion of it's translational energy is converted into internal energy. The maximum amount of energy which may be transferred from the translational to the internal modes (either electronic, vibrational or rotational) of an ion in a low energy CID process is described by the centre of mass kinetic energy E_{COM}.

$$E_{COM} = \left(\frac{E_{lab} m_t}{m_p + m_t (m_p / m_{pi})} \right)$$

where,

E_{lab} = The initial kinetic energy of the ion.

m_t = The mass of the target gas.

m_p = The mass of the projectile ion.

m_{pi} = The impact portion of the projectile ion.

The mechanism of CID depends on the time of interaction between the precursor ion and the collision gas. The most efficient energy transfer occurs when this interaction time between the ion and the target gas is similar to the period of the internal mode

which is excited. This is governed by the Massey paramter t_c/τ where $\tau = h/E$ and t_c is the interaction time between the projectile and the target. During low energy CID (typically < 200eV) the interaction time between the ion and the target gas are of the duration of 10^{-13} to 10^{-14} seconds [44,45]. This time is comparable with that of vibrational frequencies and thus this mode of excitation is predominant in low energy CID.

The much higher collision energy employed in high energy CID (typically in the keV range) results in much shorter interaction times between the ion and the target gas. These periods are comparable with the Bohr period of an electron, thus, high energy collisions are though to result in direct electron excitation of an ion. Redistribution of this energy leads to vibrational excitation of the ion followed by fragmentation in accordance with the quasi-equilibrium theory (QET) [46,47,48]. The QET postulates that the energy from excitation is randomised over the internal modes of the ion prior to fragmentation. Fragments are then produced by a series of competing consecutive unimolecular decompositions.

In low energy CID experiments the nature of the target gas plays an important role in determining the extent of fragmentation [49,50]. The use of heavier target gases such as argon increases the maximum amount of translational energy available for conversion into internal energy, thus increasing the extent of fragmentation. Target gas pressure is also another important parameter which can effect the degree of fragmentation. An increase in target pressure [51] will increase the probability of multiple collisions between any given precursor ion and the neutral target gas. This

will lead to a greater degree of fragmentation, although this may not always be desirable in every experiment as low mass/charge fragment ions are not always as informative as high mass/charge fragment ions especially when attempting to sequence proteins and peptides.

Under low energy CID conditions the efficiency of the CID process decreases significantly for singly charged ions having masses above approximately 800 Da due to insufficient energy conversion to promote fragmentation [52]. For this reason the study of singly charged ions above this mass range has historically been conducted by high energy CID. ESI has, however, significantly increased the mass of which useful CID data can be obtained by producing multiply charged ions. These ions experience a collision energy of z E_{lab} upon striking a stationary collision gas thus increasing the amount of energy available to promote fragmentations. Useful CID experiments have been performed using ESI on tandem quadrupole instruments for proteins up to several thousand Daltons, although the product ion spectra produced by the CID of highly charged precursor ions can be difficult to interpret since the number of charges carried by the fragment ions can be difficult to determine.

2.6 References

1/ A Novel hybrid quadrupole orthogonal acceleration time of flight mass spectrometer, R. H. Bateman, R. S. Bordoli, A. J. Gilbert, J. B. Hoyes, H. R. Morris, presented at ASMS 1996.

2/ J. A. Loo, G. C. Edmonds, H. R. Udseth, R. D. Smith, *Anal. Chem.*, **62**, 693, (1990).

3/ J. A. Loo, R. R. Ogorzalek-Loo, H. R. Udseth, C. G. Edmonds, R. D. Smith., *Rapid Commum. Mass Spectrom.*, **5**, 101 (1991).

4/ M. Dole, R. L. Hines, L. L. Mack, R. C. Mobley, L. D. Ferguson, M. B. Alice, *Macromolecules*, **1**, 96 (1968).

5/ M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, M. B. Alice, *J. Chem. Phys.*, **49**, 2240 (1968).

6/ G. A. Clegg and M. Dole., *Biopolymers*, **10**, 821 (1971).

7/ Lord Rayleigh, *Philos. Mag.*, **14**, 31 (1882).

8/ J. V. Iribarne and B. A. Thomspon, *J. Chem. Phys.*, **64**, 2287 (1976).

9/ B. A Thompson, J. V. Iribarne. *J. Chem., Phys.*, **71**, 11, 4451 (1979).

10/ M. Sakairi, A. L. Yergey, K. W. M. Siu, J. C. Y. LeBlanc, R. Guevremont, S. S. Berman, *Anal. Chem.*, **63**, 1488 (1991).

11/ G. Schmelzeisen-Redeker, L. Buffering, F. W. Rollgen, *Int. J. Mass Spectrom. Ion. Proc.*, **90**, 139 (1989).

12/ M. Yamashita, J. B. Fenn, *J. Phys. Chem.*, **88**, 4671 (1984).

13/ C. M. Whitehouse, R. N. Dreyer, M. Yamashita, J. B. Fenn, *Anal. Chem.*, **57**, 675 (1985).

14/ S. F. Wong, C. K. Meng, J. B. Fenn., *J. Phy. Chem.*, **92**, 546 (1988).

15/ C. K. Meng, M. Mann, J. B. Fenn, *Z. Phys. D.*, **10**, 361 (1988).

16/ G. I. Taylor, *Proc. R. Soc. London*, **A280**, 383 (1964).

17/ P. Kebarie, L. Tang, *Anal Chem.*, **65**, 27, A-972 (1993).

18/ E. Matthieu, *J. Math. Pures Appl.*, **13**, 137 (1868).

19/ J. E. Campana, *Int. J. Mass Spectrom. Ion. Phys.*, **33**, 101 (1980).

20/ J. R. Chapman, 'Practical Organic Mass Spectrometry', John Wiley and Sons, Chichester, 1985 pp. 10.

21/ N. R. Daly, *Rev, Sci. Instrum.*, **34**, 1116 (1963).

22/ A. G. Ferrige, M. J. Sedden, B. N. Green, S. A Jarvis, J. Skilling *Rapid Commun. Mass Spectrom.*, **6**, 707 (1992).

23/ A. F. Dodonov, I. V. Chernushevich, T. F. Dodonova, V. V. Raznikov, V. L. Tal'roze. *USSR Patent number 1681340A1* (Feb 1987).

24/ M. M. Sheil, R. S. Bordoli, J. Langridge., *Micromass Application note 229* (1997).

25/ W. C. Wiley, I. H. McLaren , *Rev. Sci. Instrum.*, **26**, 1150 (1995).

26/ B. A. Mamyrin, V. I. Karataev , D. V. Scmikk, V. A. Zagulin., *Sov Phys.-JETP*, **37**, 45 (1973).

27/ J. G. Boyle, C. M. Whitehouse, J. B. Fenn., *Rapid Commum. Mass Spectrom.*, **5**, 400 (1991).

28/ A. F. Dodonov, I. V. Chernushevich, V. V. Laiko., *12th Int. Mass spectrom. Conference*, Amsterdam, August 1991, Extended abstract p.153

29/ A. F. Dodonov, I. V. Chernushevich, V.V. Lakio, In 'Time of Flight Mass Spectrometry', R. .J. Cotter (Ed) American Chemical Society, Washington, DC, symposium Series 549, 1994, pp.108.

30/ O. A. Mirgorodskaya, A. A. Shevchenko, I. V. Chernushevich, A. F. Dodonov, A. I. Miroshnikov, *Anal. Chem.*, **66**, 99 (1994).

31/ P. W. Geno, In 'Mass spectrometry in the Biological Sciences: A tutorial', M. L. Gross (Ed), Nato ASI Series, Kluver Academic Publishers, 1992, pp.133.

32/ C. J. Barinaga, C. G. Edmonds, H. R. Udseth, R. D. Smith, *Rapid Commun. Mass Spectrom.*, **3**, 60 (1989).

33/ E. D. Lee, J. D. Henion, T. R. Covey, *J. Macrocolumn Sep.*, **1**, 14 (1989).

34/ T. R. Covey, E. C. Huang, J. D. Henion, *Anal. Chem.*, **63**, 1193 (1991).

35/ C. K. Meng, C. N. McEwen, B. S. Larsen, *Rapid Commun. Mass Spectrom.*, **4**, 147 (1990).

36/ V. Katta, S. K. Chowdhury, B. T. Chait, *Anal. Chem.*, **63**, 174 (1991).

37/ R. D. Voyksner, T. Pack, *Rapid Commun. Mass Spectrom.*, **5**, 263 (1991).

38/ E. de Hoffman, *J. Mass. Spectrom.*, **31**,129 (1996).

39/ K. R. Jennings, *Int. J. Mass Spectrom. Ion Phys.*, **1**, 227 (1968).

40/ R. A. Yost and C. G. Enke, *J. Am. Chem. Soc.*, **100**, 2274 (1978).

41/ R. A. Yost and C. G. Enke, *Anal. Chem.*, **51**, 1251A (1979).

42/ J. Los and T.R. Govers, In 'Collision Spectroscopy', R.G. Cooks (Ed)., Plenm Press, New York, 1978, pp. 289.

43/ K. Levsen, H. Schwarz, *Mass Spectrom. Rev.*, **2**, 77 (1983).

44/ R. N. Hayes and M. L. Gross, In 'Methods in Enzymology', **193** (Mass Spectrometry), J.A. McCloskey (Ed), 1990, pp. 237.

45/ Z. Herman, J. H. Futrell, B. Friedrich, *Int. J. Mass Spectrom. Ion. Proc.*, **58**, 181 (1984).

46/ P. J. Todd and F. W. McLafferty, In 'Tandem Mass Spectrometry', F.W. McLafferty (Ed.), John Wiley and Sons, New York, 1983, pp.149.

47/ K .Levson, 'Fundamental Aspects of Organic Mass Spectrometry' Verlag Chemie, Weinheim, New York, 1978.

48/ A. G. Brenton, R. P. Morgan, J. H. Beyon, *Annv. Rev. Phys. Chem.*, **30**, 51 (1979).

49/ R. P. Grese, M. L. Gross, *J. Am. Chem. Soc.*, **112**, 5098 (1990).

50/ L. M. Teesch, J.Adams, *J. Am. Chem. Soc.*, **113**, 812 (1991).

51/ V. H. Wysocki, H.I. Kenttamaa, R.G. Cooks, *Int. J. Mass Spectrom. Ion. Proc.*, **75**, 181 (1987).

52/ A. J. Alexander, R. K. Boyd, *Rapid Commun. Mass Spectrom.*, **3**, 211 (1989).

Chapter 3

Characterisation Of The Alkene Monooxygenase From *Norcardia corallina* B276

3.1 Introduction

The bacteria *Nocardia corallina* B-276 expresses a multicomponent enzyme called the alkene monooxygenase (AMO).

The AMO has the ability to convert alkenes to optically active epoxides [1]. Epoxides contain a three membered heterocyclic ring. Substitution on one of the carbon atoms results in the formation of a chiral centre and hence produces an optically active compound. Due to the high reactivity of epoxides they have been used in a large number of organic synthetic routes to chemically and biologically important compounds. Specific uses of epoxides include the synthesis of prostaglandins, the development of ferro-electric liquid crystals [2] and in the synthesis of S-arylgycidyl ethers used in the production of β-blockers [3].

Traditional synthetic methods of producing epoxides include the Sharpless catalyst [4] and the use of metalloporphyrins [5]. These methods, however, have several disadvantages; for this reason attention has turned to biological methods of epoxygenation.

In general, monooxygenases catalyse the direct insertion of one atom of oxygen into an organic compound using NADH (reduced form of nicotinamide adenine dinucleotide) as the electron donor. Monooxygenases from many micro-organisms have been purified and studied. Only a few organisms, however, that grow on simple n-alkenes and n-

alkanes have been considered. Probably the most studied alkane-using enzyme is the soluble methane monooxygenase (sMMO) [6], which catalyses not only hydroxylation of methane but also oxidation of many kinds of compounds such as alkanes, alkenes, ethers, aromatic hydrocarbons and even carbon monoxide and ammonia.

MMO has been purified from several methanotrophs and consists of three components: a hydroxylase, a NADH acceptor reductase and a regulatory protein.
Similarly the AMO enzyme consists of three components all of which are considered to be necessary for activity; an epoxygenase, an NADH dependent reductase and a coupling protein [7].

The epoxygenase consists of two subunits of RMM 57.1 and 38.3 kDa respectively and is believed to be the site of alkene epoxygenation. The second component called the reductase is a 37.2 kDa protein which contains a FAD group and an Fe_2S_2. The role of the final component, termed the coupling protein RMM of 12.8 kDa, is undetermined but is essential for activity of the AMO complex.

3.2 Experimental

Mass Spectrometry

All mass spectra were recorded by means of a "Quattro II" QhQ tandem mass spectrometer (Micromass UK Ltd, Altrincham, UK) equipped with an atmospheric pressure ionisation source operated in the nebulizer-assisted electrospray mode. The

potential of the electrospray needle was set at 3.5 kV and the extraction cone voltage was varied linearly from 30V at m/z 600 to 90V at m/z 2600. The source temperature was held at 60°C. The carrier solvent used was a 1:1 mixture of CH_3CN/H_2O containing 0.1% HCOOH. The samples were infused using a syringe pump (Harvard model 11, South Natic, USA) at a flow rate of 5μL/min. Calibration was carried out using a solution of horse heart myoglobin or sodium iodide. Mass spectra were acquired over the range m/z 600-2000 during a 10s scan and by operating the data system in the multichannel acquisition (MCA) mode.

For LC/MS analysis a Perkin Elmer Model 140 A system equipped with a C_8 reverse phase HPLC column (Brownlee, 250 x 1 mm) was coupled to the mass spectrometer. The flow rate of 50 μL/min was split after the column such that about 20 μL/min was directed into the mass spectrometer. Native epoxygenase analysis was carried out using a linear gradient from 20 to 90 % of solvent B over 50 min. The spectra were acquired from m/z 600-1800 for the hydroxylase analyses and by operating the data system in the continuum mode. Solvent B consisted of 90% CH_3CN plus 0.1% trifluoroacetic acid.

All data were processed by means of the MassLynx data system. The reconstructed mass spectra were obtained by using the Maximum Entropy algorithm.

3.3 Results and discussion

3.3 (a) The epoxygenase

The ESI spectrum of the epoxygenase is shown in figure 3.1. The spectrum shows a series of peaks corresponding to multiply charged ions ranging from +22 to +48 charges. The ions correspond to a protein of RMM 38332 Da. This protein is assigned as the smaller subunit of the epoxygenase. A second series of multiply charged ions is also visible in the spectrum, although this is far less prominent than the first.

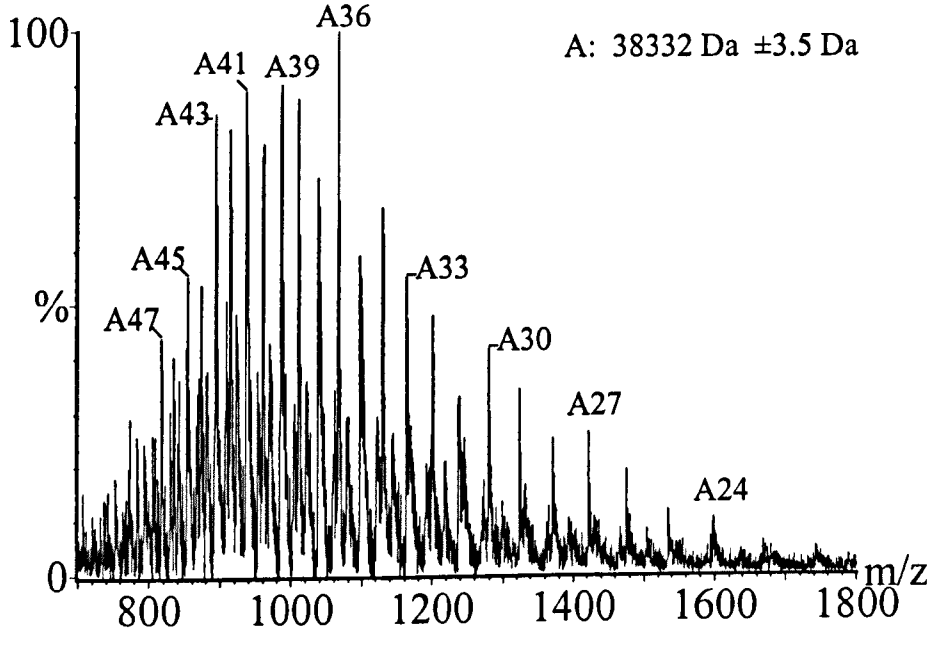


Figure 3.1. ESI spectrum of the epoxygenase

The presence of more than one protein was more readily demonstrated by the "Maximum Entropy" ("MaxEnt") deconvoluted mass spectrum. This clearly shows the presence of the protein at 38331 Da already assigned as the smaller subunit of epoxygenase and another protein of RMM 57081 Da. This protein is assigned as the larger subunit of the epoxygenase. The "MaxEnt" deconvoluted mass spectrum of the epoxygenase is shown in Figure 3.2.

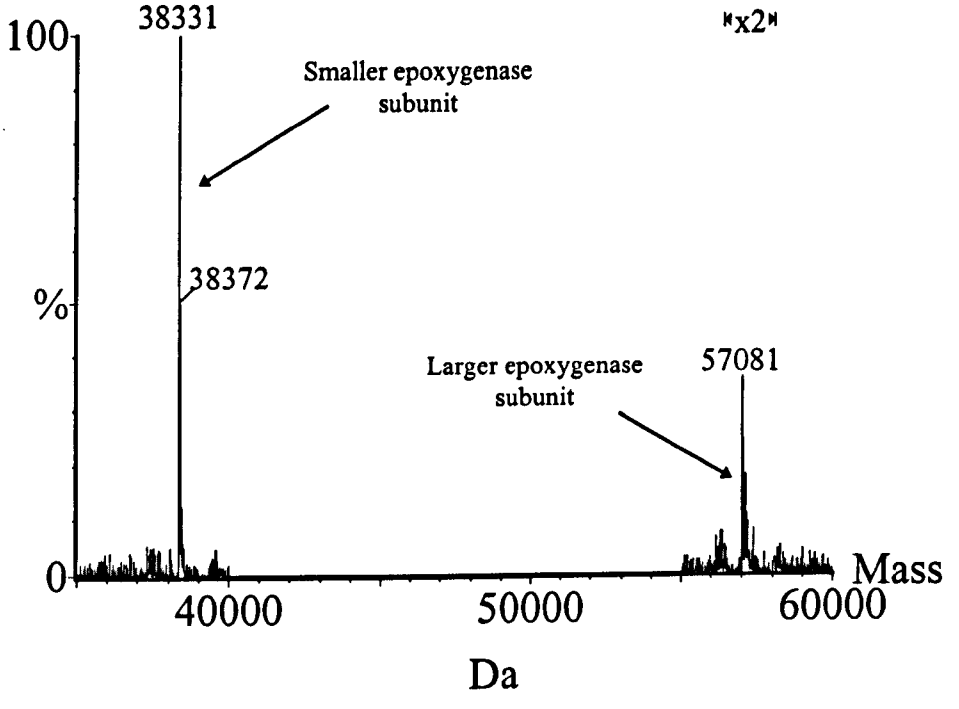


Figure 3.2. "MaxEnt" deconvoluted mass spectrum of the epoxygenase

Improved detection of the larger subunit of the epoxygenase was achieved by submitting the epoxygenase to HPLC-ESI-MS.

The total ion current (TIC) profile of the epoxygenase is shown in figure 3.3. The two expected subunits of the epoxygenase were located in the main peaks of the TIC profile (retention times 30.2 and 36.0 minutes respectively). The minor peaks on the TIC profile were found to consist of impurities.

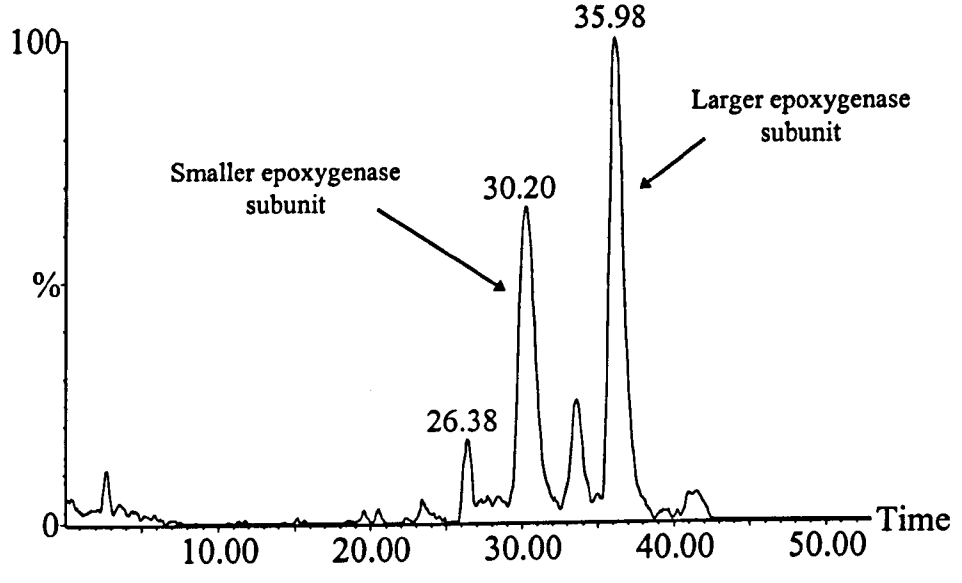


Figure 3.3. Total ion current (TIC) profile of the epoxygenase

The spectrum which produced the peak that eluted at 30.2 minutes of the TIC profile is shown in figure 3.4.

The spectrum shows a series of peaks corresponding to multiply charged ions of a protein of RMM 38331 Da, this protein is assigned as the smaller subunit of the

epoxygenase as identified in the direct injection of the epoxygenase as shown in figure 3.1.

The measured RMM of the smaller subunit is 38331 Da. A comparison of this value with the predicted RMM of 38462 Da, as calculated from gene sequencing, suggests that the N-terminal methionine residue of the polypeptide have been cleaved. If one takes this process into account the predicted RMM of the smaller subunit is 38331 Da. This value is in excellent agreement with the measured RMM, and suggests that the published primary sequence of the smaller subunit of the epoxygenase is correct

The ESI spectrum that produced the peak eluted at 36.0 minutes of the TIC profile of the epoxygenase is shown in figure 3.5.

The spectrum shows a series of peaks which correspond to multiply charged ions (32 to 74 charges) of a protein of measured RMM 57080 Da. This protein is assigned as the larger subunit of the epoxygenase.

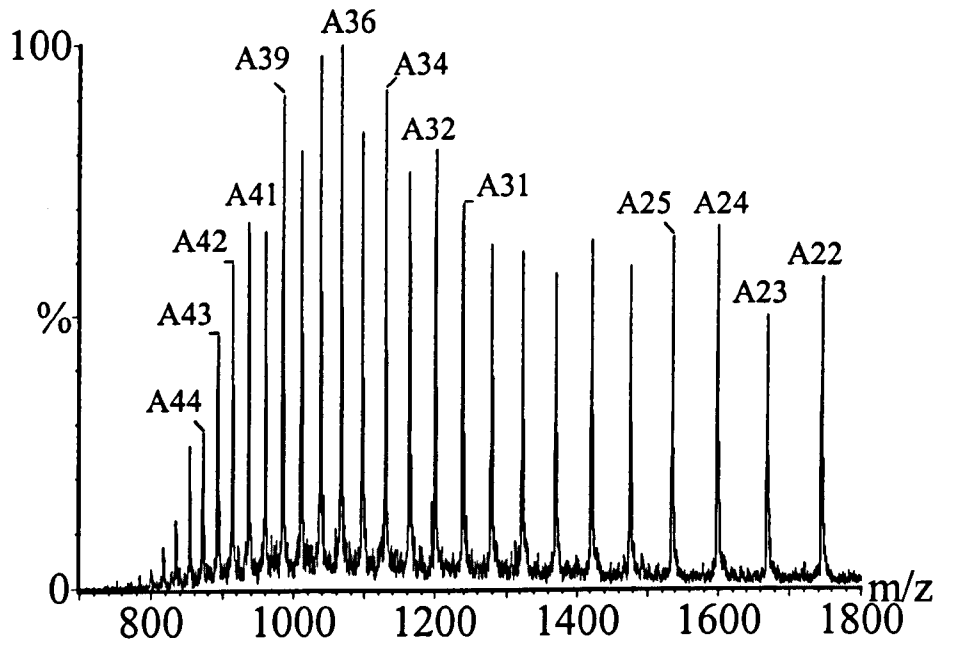


Figure 3.4. The ESI spectrum that produced the peak eluted at 30.2 minutes of the TIC profile of the epoxygenase

Again comparison with the predicted RMM of 57211 Da, as calculated by gene sequencing, suggests that the N-terminal methionine of the polypeptide has been cleaved. This result suggests that the published primary sequence of the larger epoxygenase subunit is also correct.

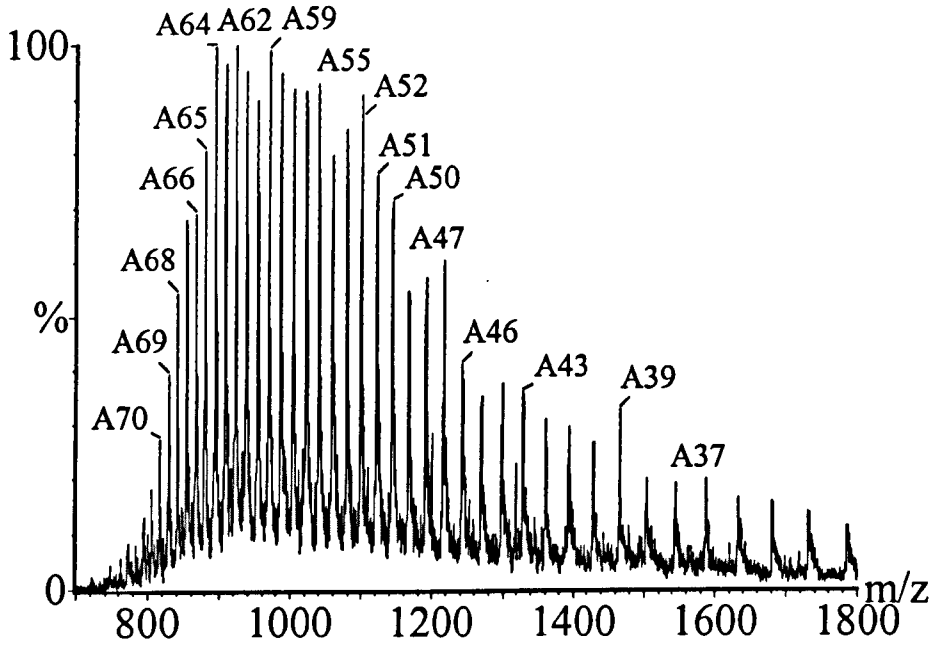


Figure 3.5. ESI spectrum that produced the peak eluted at 36.0 minutes of the TIC profile of the epoxygenase.

3.3 (b) The Reductase

The single component reductase was analysed by direct injection into the mass spectrometer. The ESI spectrum of the reductase is shown in Figure 3.6.

The spectrum shows a series of peaks which correspond to multiply charged ions of the charge states +23 to +48. The ions correspond to a protein of RMM 37152 Da. The single component reductase is easier to visualise in the "MaxEnt" deconvoluted mass spectrum as shown in figure 3.7.

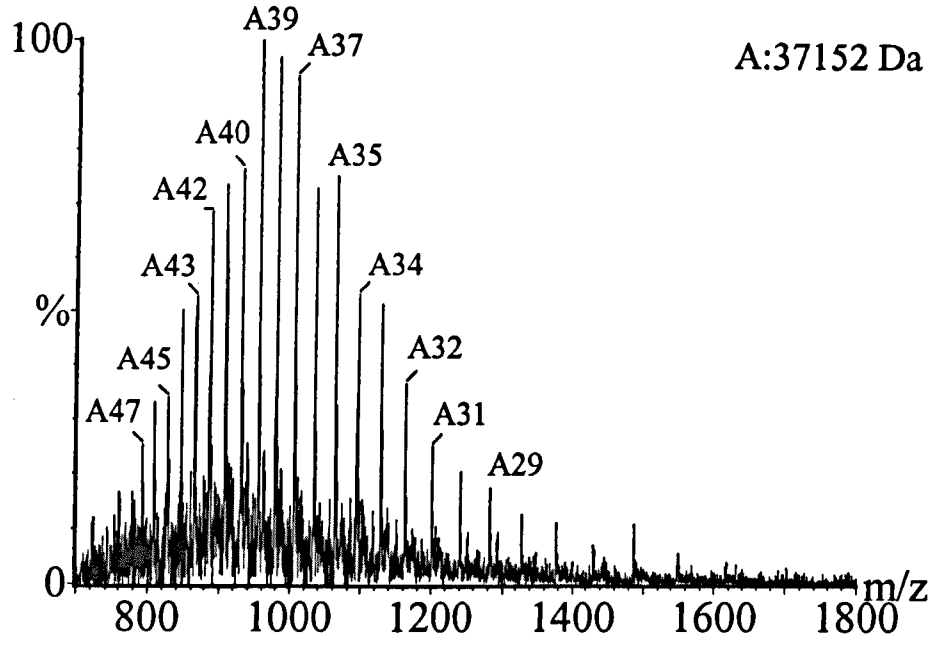


Figure 3.6. ESI spectrum of the reductase

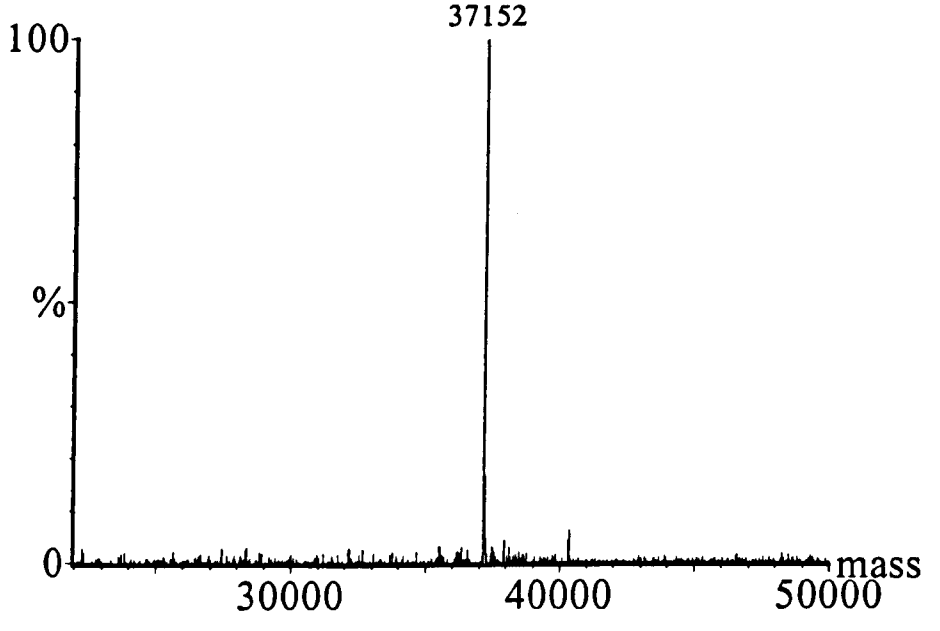


Figure 3.7. "MaxEnt" deconvoluted mass spectrum of the reductase.

The predicted RMM, as calculated from gene sequencing, is 37282 Da. If removal of the N-terminal methionine occurs giving a RMM of 37151 Da, then this value is in good agreement with the measured RMM of 37152 Da. These data suggest that the published amino acid residue sequence of the reductase is correct, although its predicted RMM reported by Saeki and Furuhasi [8] of 37828 Da appears to have been miscalculated.

3.3 (c) The coupling protein

The ESI spectrum of the coupling protein is shown in Figure 3.8. The protein is prepared in the presence of benzamidine, a serine protease inhibitor, since in its absence, it has been observed that cleavage of the coupling protein occurs resulting in loss of all AMO activity [9].

A similar cleavage of protein B, the equivalent protein from the sMMO complex has been observed during its purification which also results in the loss of activity of the enzyme.

The ESI spectrum of the coupling protein, prepared in the presence of benzamidine, shows the presence of more than one series of multiply charged ions. This suggests that benzamidine does not completely stop the cleavage of the coupling protein.

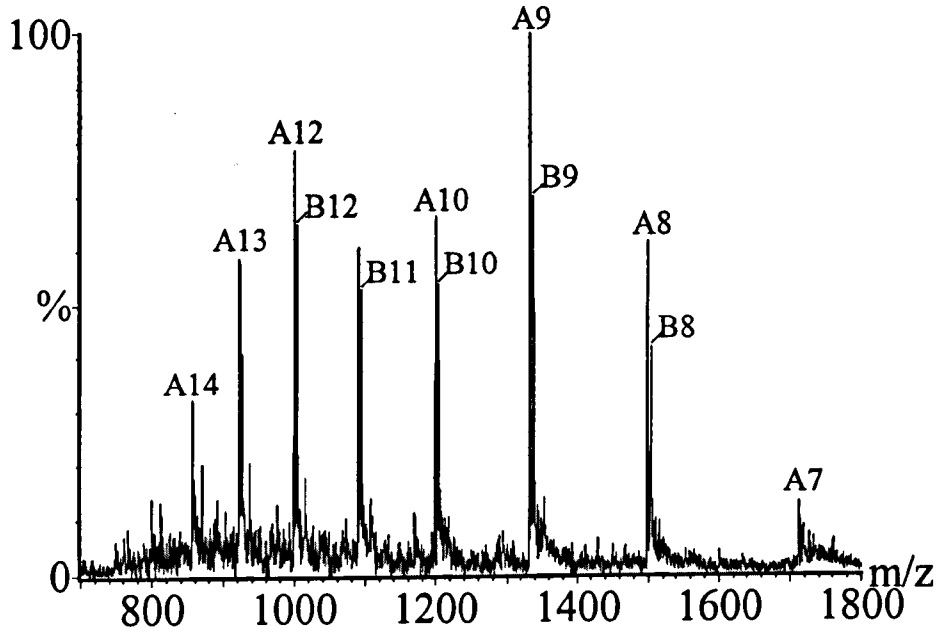


Figure 3.8. ESI spectrum of the coupling protein prepared in the presence of benzamidine

The presence of several proteins is demonstrated in the "MaxEnt" deconvoluted mass spectrum of the coupling protein as shown in Figure 3.9.

The coupling protein has a predicted RMM of 12985 Da, as calculated from gene sequencing. Removal of the N-terminal methionine, as seen for the other subunits of the AMO, would result in a predicted RMM of 12854 Da. This value is higher than any of the proteins detected in the spectrum of the coupling protein.

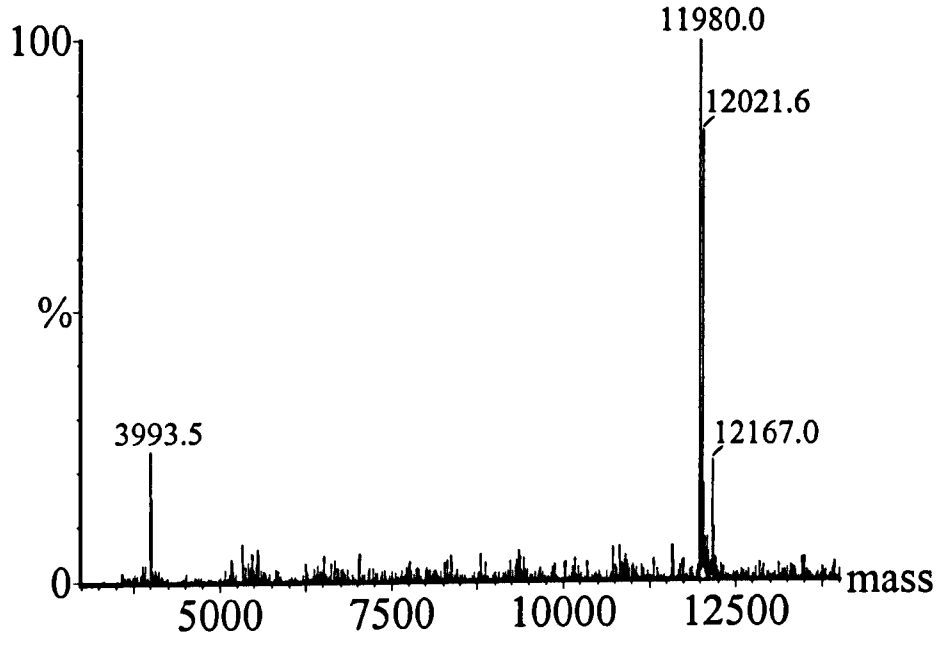


Figure 3.9. "MaxEnt" deconvoluted mass spectrum of the coupling protein prepared in the presence of benzamidine

The most prominent proteins detected by ESI-MS had RMM values of 11980.0, 12021.6 and 12167.0 Da respectively. If the predicted amino acid sequence is correct the 12167 Da protein is most likely to have arisen from a cleavage between Lys4 and Glu5 and residues Phe114 and Lys115. The 11980 Da is most likely due to a single cleavage between residues Val8 and Thr9. There are two possible cleavage reactions which could have resulted in the 12021.6 Da protein as illustrated in Table 3.1.

Measured RMM (Da)	Residues	Predicted RMM (Da)	Difference
12167.1	5-114	12167.6	0.5
12021.6	3-111	12020.4	-1.2
12021.6	5-113	12020.4	-1.2
11980.0	9-116	11980.4	0.4

Table 3.1. Measured RMM values and possible cleavage sites in the proteins present in the preparation mixture containing benzamidine.

Purification of the coupling protein in the absence of benzamidine resulted in the production of an inactive protein (in terms of AMO epoxygenation activity). The ESI spectrum of this preparation is shown in figure 3.10.

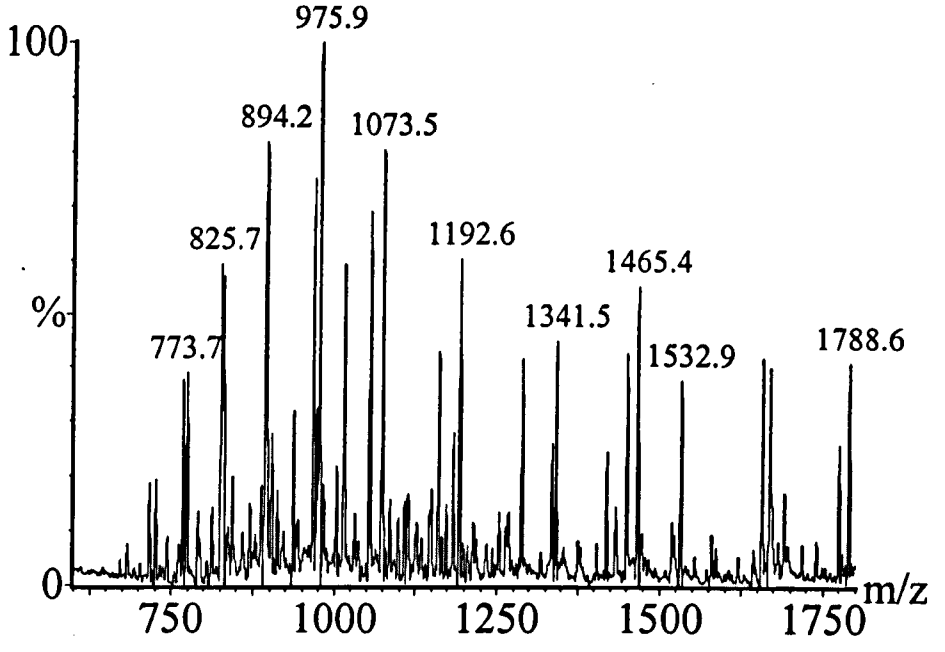


Figure 3.10. ESI spectrum of the coupling protein prepared in the absence of the serine protease inhibitor benzamidine

The ESI spectrum of the coupling protein prepared in the absence of benzamidine shows a complicated mixture of multiply charged ions, suggesting that the coupling protein has undergone further cleavage when prepared without benzamidine. The "MaxEnt" deconvoluted mass spectrum shows the presence of several cleavage proteins as shown in figure 3.11.

The most prominent proteins in the spectrum are 12627.5 Da, 12167 Da, 11590.0 Da and 10723.0 Da. The 12167.0 Da species is probably the same species observed as the

material purified in the presence of benzamidine. The 11590 Da peptide was not observed when the protein was purified with benzamidine.

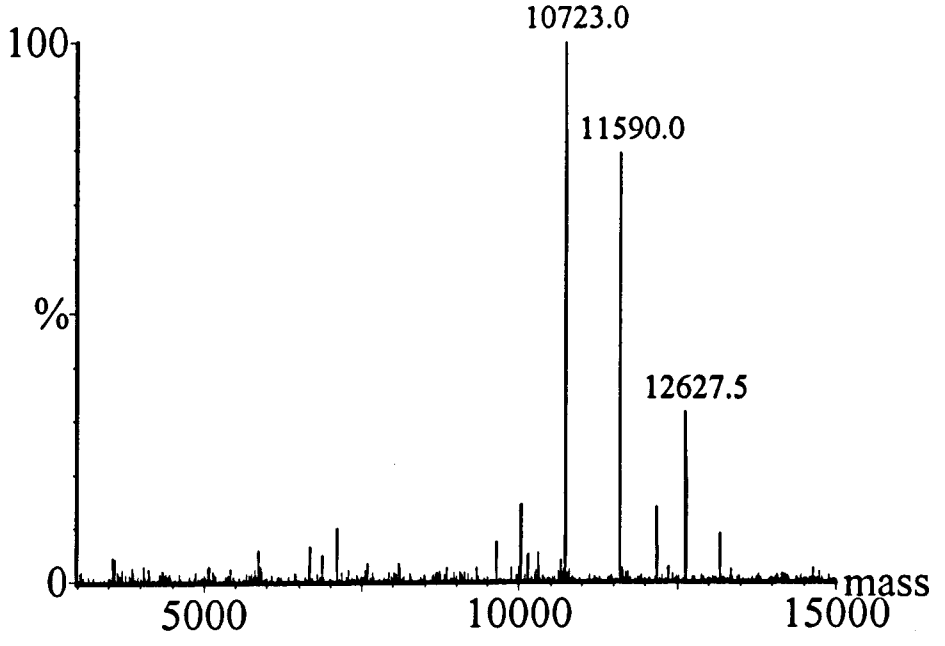


Figure 3.11. "MaxEnt" deconvoluted mass spectrum of the coupling protein prepared in the absence of benzamidine.

Analysis of the amino acid sequence of the coupling protein indicates that the 11590 Da peptide may have arisen from a cleavage between residues Arg113 and Phe114. The likely cleavages of the coupling protein which would result in the 10723.0 Da species are illustrated in Table 3.2.

The 12627.0 Da species could not be identified as a cleavage product and could represent an impurity that co-purified with the coupling protein.

It is most likely that the 11980 Da and/or the 12021.6 Da species are the active form/s of the protein when purified in the presence of benzamidine since the 12167 Da peptide was present whether benzamidine was added or not.

If the 11980 Da species is the active form of the coupling protein this suggests that possibility the C-terminus of the coupling protein is essential for the activity of the coupling protein as this is retained in this cleavage protein. There are two possible consequences for the cleavage of the coupling protein at the N-terminus. First, that the N-terminus is simply not required for activation of the protein, or that cleavage is required for activation of the protein in a similar way to that observed for the activation of chymotrypsinogen to chymotrypsin.

Alternatively the 12021.6 Da cleavage protein could be the active form of the coupling protein as this was only observed when the coupling protein was prepared in the presence of benzamidine.

Measured RMM (Da)	Residues	Predicted RMM (Da)	Difference (Da)
12627.8	???	N/A	N/A
12167.8	5-114	12167.6	-0.2
11590.2	9-113	11590.0	-0.2
10723.1	10-106	10723.0	-0.1
10019.8	7-97	10019.2	-0.6
10019.8	10-100	10019.2	-0.6
7091.0	49-109	7090.0	-1.0
7091.0	50-110	7090.0	-1.0
6668.6	46-103	6668.5	-0.1

Table 3.2. Measured RMM values and possible cleavage sites of the proteins present in the preparation mixture when benzamidine was not present.

3.4 Conclusions

The good agreement between the measured and predicted RMM values for each of the component proteins of the AMO suggests that the published amino acid sequences are correct.

All the components have at some stage after synthesis been processed by the removal of the N-terminal methionine.

Analysis of the coupling protein has identified cleavage reactions which lead to inactivation of the protein. The inclusion of benzamidine during the purification of the coupling protein prevents cleavage at the C-terminus in the production of the 11980 Da protein which could prevent inactivation of the coupling protein.
A similar cleavage reaction was observed at the N-terminus of protein B of sMMO which resulted in inactivation of the enzyme. Cleavage of protein B could not be prevented by addition of a protease inhibitor to the preparation
Alternatively benzamidine could prevent further cleavage of the 12021.6 Da protein which could be the active form of the protein.

3.5 References

(1) K. Furuhashi, A. Taoka, S. Uchida, I. Karube, S. Suzuki, *Eur. J. Appl. Microbiol. Biotechnol.*, **12**, 39 (1981).

(2) K. Furuhashi, In 'Charality in industry', A. Collins, G. Sheldrake and J. Crosby (Eds.), John Wiley and sons, London, pp167.

(3) D. Leak, P. Aikens, M. Seyed-Mahmoudian, *Tibtech.*, **10**, 256 (1992).

(4) V. Martin, S. Woodard, T. Katsuki, Y. Yamada, M. Ikeda, K. Sharpless, *J. Am. Chem. Soc.*, **103**, 6237 (1981).

(5) Z. Wei, J. L. Loebach, S. R. Wilson, E. N. Jacobsen, *J. Am. Chem. Soc.*, **112**, 2801 (1990).

(6) H. Dalton, In 'Methane and Methanol Utilizers' J. C. Murrel and H. Dalton (Eds.), Penum Press, New York, 1992, pp. 85.

(7) A. Miura and H. Dalton, *Biosci. Biotech. Biochem.*, **59**, 853 (1995).

(8) H. Saeki and K. Furuhashi, *J. Fement. Biol.*, **78**, 399 (1994).

(9) A. Miura, M.S.c. Thesis University of Warwick (1993).

Chapter 4

Characterisation Of The Ceroid Lipofuscinosis Protein; A Model System For Subunit c Of The Mitochondrial ATP Synthase

4.1 Introduction

The neuronal ceroid lipofuscinoses, known collectively as Batten disease, are a group of neurodegenerative lysosomal storage diseases affecting humans and other animals [1]. Three main forms of the human disease exist and are classified by the age of the victim upon onset of the disease. These are; an infantile form (INCL or CL1) which occurs most frequently in Finland, a late infantile form (LINC or CLN2) and a juvenile form (JNCL or CLN3) [2]. Children affected with the disease lead a normal life until at different ages they develop increasing dementia, blindness and seizures culminating in premature death. Collectively these diseases have been estimated to occur at a rate of one in every 12,500 births world wide [3].

The various forms of Batten diseases are proteinoses in which specific hydrophobic proteins are stored in lysosome-derived organelles. This initial discovery was made in a well described form of Batten disease in a flock of sheep deliberately bred and maintained as a model of the human disease. Analysis of the storage bodies isolated from fresh tissue revealed that more than 70% of their mass was protein, the rest being lipids typical of lysosome-derived organelles [4].

Initially the major storage body was difficult to identify as it is insoluble in all but a few solvents. Direct N-terminal sequencing eventually revealed that the major storage body protein had the same complete sequence as subunit c of mitochondrial ATP (adenosine triphoshate) synthase [5].

These results established that the sheep disease, ovine ceroid lipofuscinosis, is a proteinosis in which subunit c of ATP synthase, an hydrophobic inner mitochondrial membrane protein is abnormally stored in lysosome derived organelles. Subsequent studies showed that a protein with an identical amino acid residue sequence to that of complete subunit c is also stored in the late infantile and juvenile human diseases [6]. N-terminal sequencing also established subunit c storage in affected English setter, Border collie and Tibetan terrier dogs [7]. The relative amounts of subunit c in storage bodies varies in these diseases and there is evidence of co-storage of other proteins in some forms.

The metabolic basis for the accumulation of subunit c in the ceroid lipofuscinoses is currently unknown. The specificity of the storage suggests that the concept underlying Batten disease are lesions that effect the turnover of subunit c in such a manner that it accumulates in storage bodies [6]. Possibilities include, the over production of subunit c or a modification of subunit c which is specific to the disease and causes accumulation. In particular Katz and Rodrigues [8] have suggested that the stored subunit c carries a trimethyl-modification of one of its two lysine residue (Lys 43), although this finding has been questioned by others [9].

Subunit c is found in mitochondrial ATP synthase (F_1-F_0-ATPase) which is a mutli-subunit enzyme complex that utilises a transmembrane proton gradient to synthesise ATP [10]. It consists of two domains linked by a stalk assembly (called the γ-subunit) : a hydrophilic part, F_1, which is the catalytic site of ATP synthesis and a

transmembrane portion, F_0, which catalyses proton translocation events (Figure 4.1.). The number of F_0 subunits is known to be variable and species specific. Subunit c is known to be present in multiple copies, in the F_0 subunit, which are proposed to form the components of an oligomeric proton channel. The mechanism of proton conduction is not known but a major role has been attributed to Glu-58 in mitochondrial subunit c.

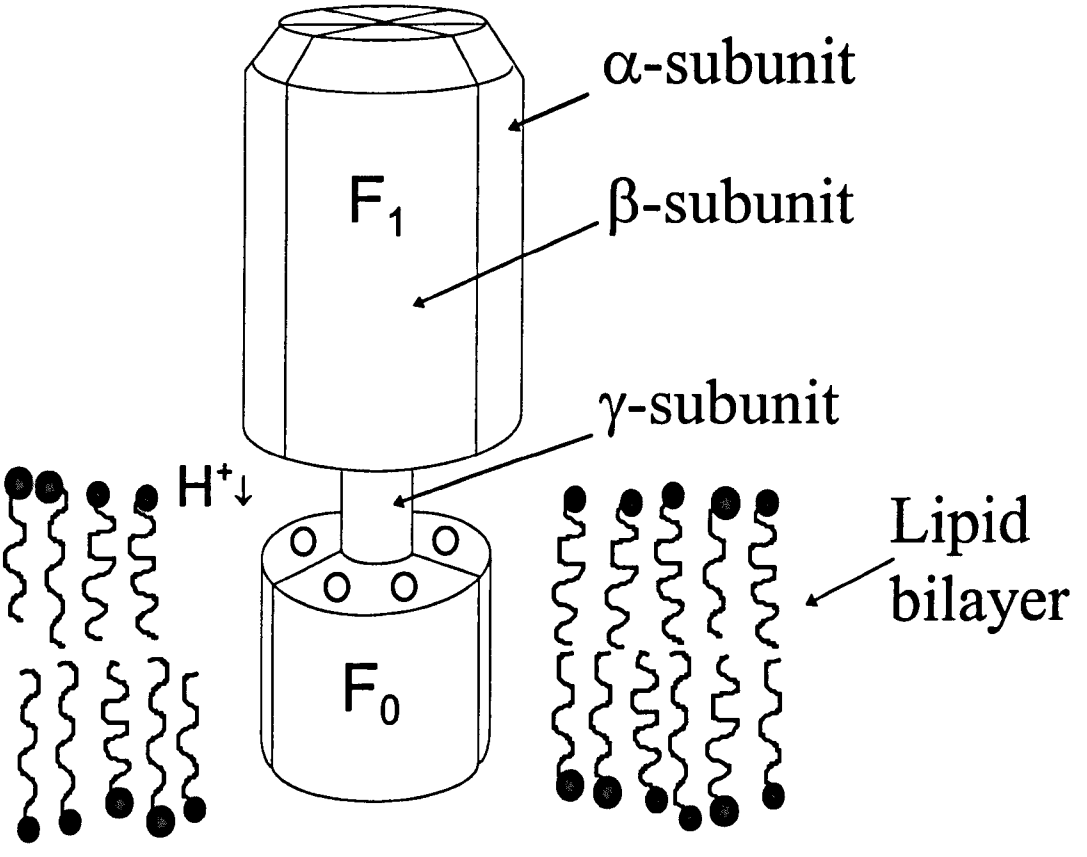
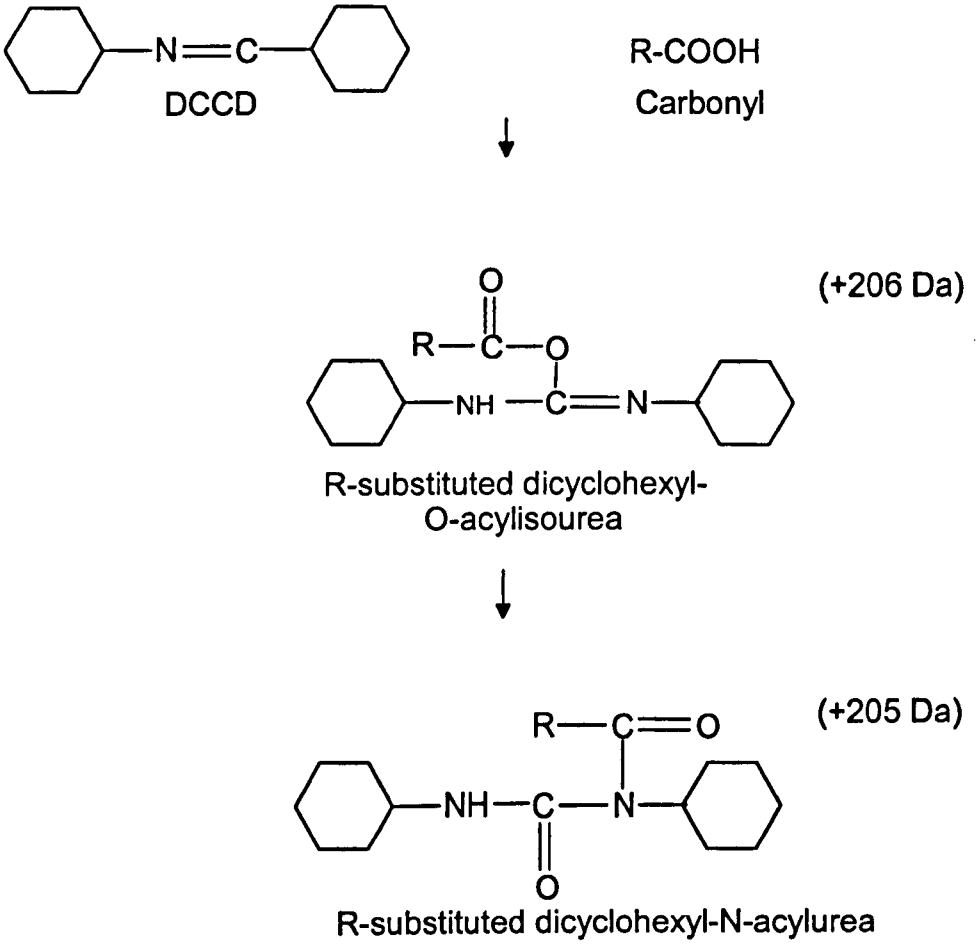


Figure 4.1. Schematic representation of the Mitochondrial ATP synthase (F_1-F_0-ATPase) spanning the inter mitochondrial membrane.

The subunit c protein is also known to be the site of binding of several diagnostic inhibitors of the ATP synthase such as N,N'-dicyclohexylcarbodiimide (DCCD). DCCD is thought to block ATP synthesis by interfering with the utilisation of the proton gradient.



Scheme 4.1. Schematic of the intramolecular rearrangement of an R-substituted dicyclohexyl-O-acylisourea to an R-substituted dicyclohexyl-N-acylurea

DCCD interacts with a specific β-carboxyl group in the F_O-subunit [11]. The reaction product is an unstable O-acylisourea (+206 Da) which by intramolecular rearrangement yields a stable N-acylurea with a final mass increase of 205 Da. [12] as illustrated in Scheme 4.1.. Studies of the correlation of [^{14}C]-DCCD binding with heart mitochondria show that the [^{14}C]-DCCD binding with the equivalent of only one subunit c out of the circular array of 9-12 subunit c present in F_O correlates with maximal inhibition of ATP synthase activity [13]. There is, however, a requirement for at least an additional 2-3 moles of DCCD per mole of F_1F_0 ATP synthase for incorporation of 1 mole of [^{14}C]-DCCD per mole F_1F_0 ATP synthase. The function of the excess DCCD is unknown but is usually attributed to non-specific hydrolytic or non-productive condensation reactions which are not inhibitory.

Subunit c is thought to fold in the Fo sector of the ATP synthase like a hairpin with two membrane traversing α-helices as shown in Figure 4.2 [14]. Residues in both membrane traversing helices are labelled by hydrophobic agents that are thought to react with the apolar phase of the lipid bilayer. Proton transport through the membrane along F_0 probably induces long-range conformational changes that pass via the stalk to the catalytic sites in F_1 to drive ATP synthesis.

Mitochondrial subunit c is both difficult to extract and purify due to its hydrophobic nature. Alternatively CLP can be extracted in high yield and purity from the lysosomal storage bodies of sheep affected with Batten disease [5]. CLP has significant advantages

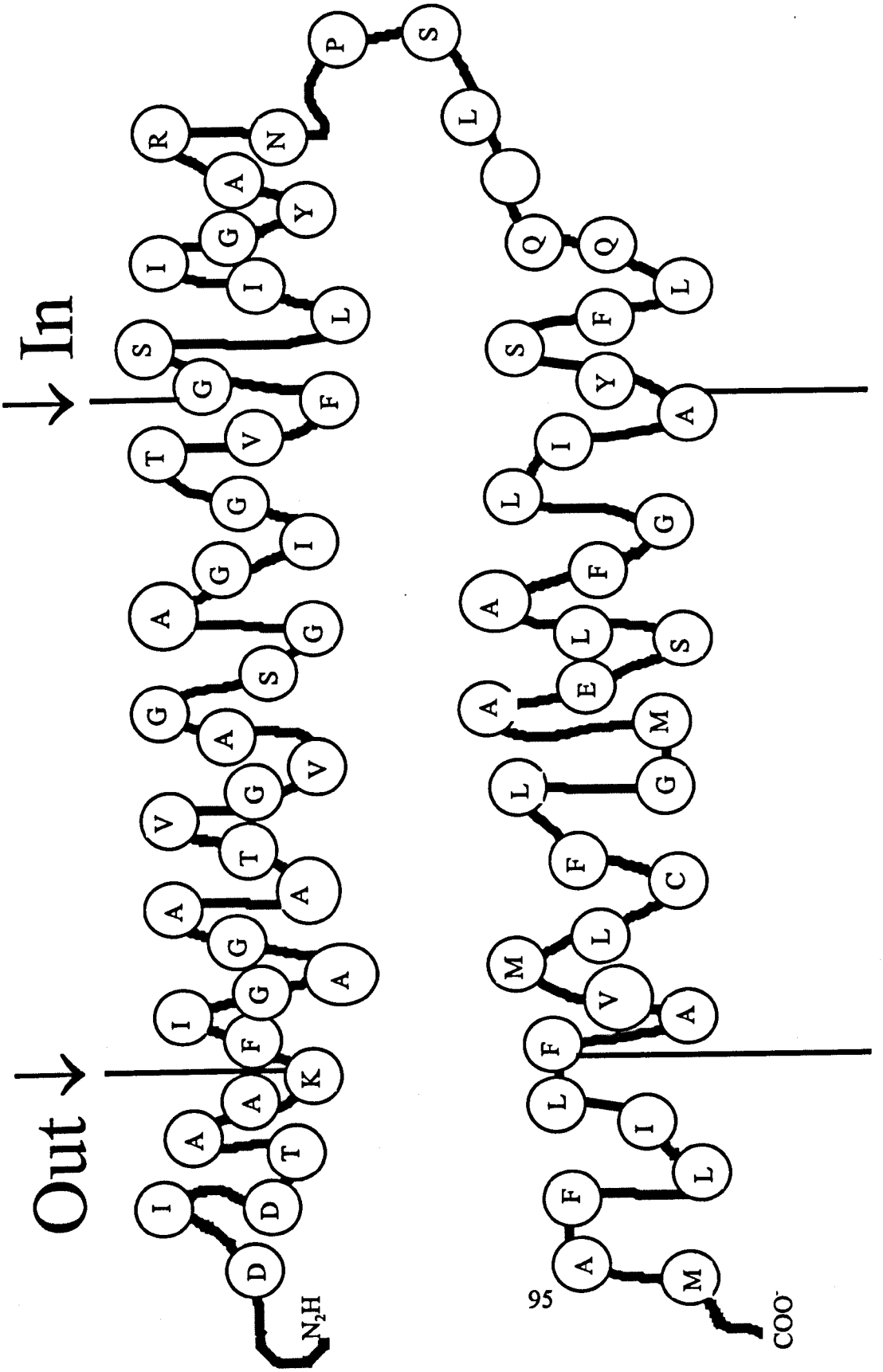


Figure 4.2. Folding of mitochondrial subunit c in the coupling membrane.
Adapted from Hassinen & Vuolika, *Biochem. Biophys.Acta.*, **1144,** 107 (1993)

therefore as a model system of mitochondrial subunit c and its interactions with inhibitors for ESI-MS studies [15].

In order to span the inner mitochondrial membrane subunit c is, as mentioned previously, hydrophobic. Hydrophobic proteins are not routinely characterised by ESI-MS as they are generally insoluble in the solvents used in the ESI process. ESI is also intolerant of detergents which are used quite often to solubilize such proteins. Some hydrophobic proteins can, however, be characterised by ESI-MS when dissolved in suitable solvents such as chloroform/methanol/water (C/M/W) [16]. This solvent system is used in this study both in the extraction of the CLP and subunit c and as the solvent system used for the ESI.

4.2 Experimental

Mass spectrometry

All m/z spectra were acquired by means of a "Quattro II" tandem mass spectrometer of QhQ geometry operated in the positive ion mode. The instrumental conditions were as described in chapter 3. The CLP was diluted to approximately 2pm/mL in chloroform/methanol/water (C/M/W) before ESI analysis.

Purification of the ceroid lipofuscinosis protein

Lysosomal storage bodies were isolated from the liver of sheep affected with ceroid lipofuscinosis as described by Palmer *et al* [17]. Purified CLP was extracted from the bodies by chloroform/methanol extraction [5].

Inhibitors and incubation conditions

CLP solutions were diluted with an equal volume of C/M/W and DCCD added to give a 100:1 molar ratio. The solutions were then incubated at room temperature for 90 mins or 18 hours. Samples of the incubation mixture were then further diluted with C/M/W, centrifuged at 14,000g for 1 min and immediately submitted for ESI-MS analysis. CLP solutions of differing pH were made by mixing CLP solutions containing ammonium acetate at pH 7-7.5 with trifluoroacetic acid (TFA) in the C/M/W dilution mixture at levels of 0.1% v/v TFA or 1.0% TFA, before incubation with the DCCD.

Bovine heart mitochondrial subunit c extraction

A/ Standard procedure

1ml of beef heart mitochondrial suspension (approximately 25 mg) was extracted with 25 ml C/M by stirring at 4°C overnight. The extraction mixture was filtered twice through a glass fibre filter (Whatman GF/C) using a glass filtration system (Millipore). The single phase filtrate was washed with 0.2 vol. water by vigorous shaking for 5mins and the emulsion allowed to stand for phase separation and then centrifuged at 4000g for 15 mins. The aqueous layer was removed by aspiration together with any lipid layer present at the interface of the solvent layers.

The lower phase was then washed with C/M/W (1:1v/v), using one third the volume used in the first wash. The lower layer was then used for ESI-MS analysis.

B/ Rapid micro-extraction method

A procedure was devised which is suitable for small samples (1-3 mg of protein) using micro centrifuge tubes and a bench micro-centrifuge. 0.1 mL heart mitochondrial suspension was extracted with 1.5 mL C/M by vigorous shaking for 20 min at room temperature in a 2.0 mL polyethylene micro-centrifuge tube. The resultant suspension was allowed to stand for 1 min and then centrifuged at 20,000 g for 3 min at room temperature. The floating mitochondrial layer present in a small aqueous surface layer was displaced to the side of the tube by a small spatula (metal or polyethylene) and the extract (approximately 1.5 mL) was poured directly into another micro-centrifuge tube. Water (0.3 mL) was added and after vigorous mixing for 1 min the mixture was centrifuged at 20,000g. The upper aqueous layer and any interface film is removed carefully by aspiration. The lower chloroform layer was then washed with 0.1 mL methanol / water (1:1 v/v) as before and the lower chloroform layer obtained after centrifugation was used directly for ESI-MS analysis. The mitochondrial floating layer left after the initial extraction can be re-extracted by immediate addition of 0.05 mL water followed by 0.75 mL of C/M.

4.3 Results and discussion

4.3 (a) ceroid lipofuscinosis protein

Figure 4.3. shows the ESI spectrum of the liver ceroid lipofuscinosis protein (CLP). The spectrum shows peaks corresponding to the 3+ to the 8+ ions of a protein of measured RMM 7650.2 Da ± 0.3 Da, this protein is assigned as the CLP. The measured RMM value disagrees with the predicated RMM value of 7608.0 Da, calculated by means of gene sequencing [18], by +42.2 Da. The increase in RMM could be explained by a trimethylation (+43 Da) or by an acetylation (+42 Da) of an amino acid residue. Other explanations for discrepancies in RMM include amino acid residue substitutions such as; Gly - Val (42 Da), Leu - Arg (42 Da) and Ile - Arg (43 Da).

The spectrum also shows another series of peaks which correspond to multiply charged ions of a protein of RMM 7666.2 Da. This is assigned as being oxidation of one of the amino acid residues, such as one of the methionine residues [19] present in the CLP. All attempts to follow the standard procedure of cleaving or digesting the protein in order to produce a cleavage/digest peptide map to aid in the location and identification of the RMM discrepancy were unsuccessful. This was due to the protein being hydrophobic and thus insoluble in the solvent systems employed for chemical cleavage / enzymatic digestions.

An MS/MS experiment of the intact subunit c was therefore undertaken to try and locate and identify the discrepancy in RMM.

The MS/MS spectrum of the $[M+5H]^{5+}$ ion of the CLP protein is shown in Figure 4.4. The most prominent fragment ions are a series of quadruply charged b and singly charged y ions as labelled in Figure 4.4. The singly charged y ions correspond to the y_1-y_{10} fragment ions which originate from the C-terminus. The fragment ions y_1-y_6 are accompanied by a satellite ions of +16 Da.

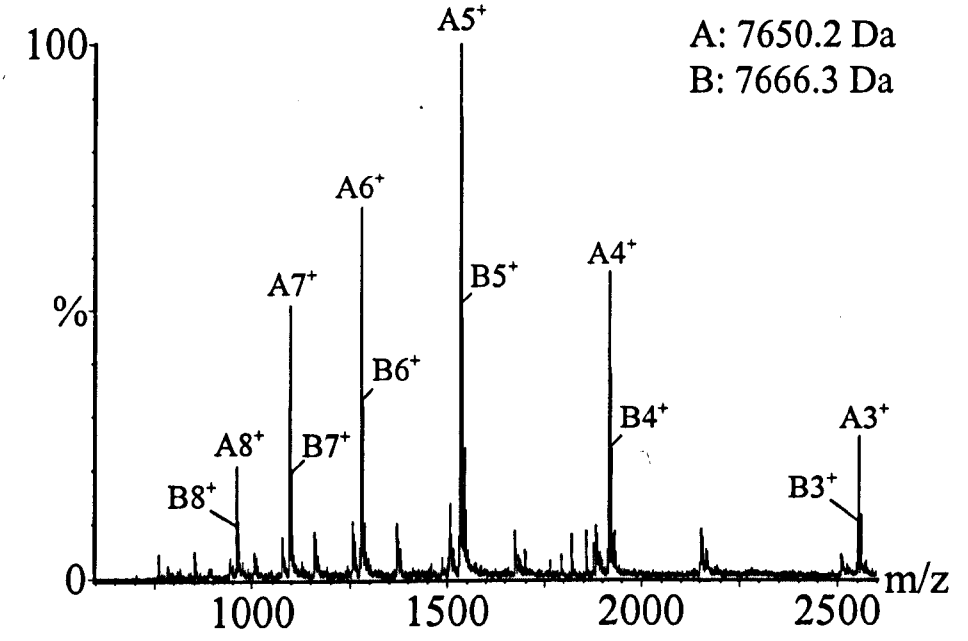


Figure 4.3. ESI spectrum of the liver ceroid lipofuscinosis protein.

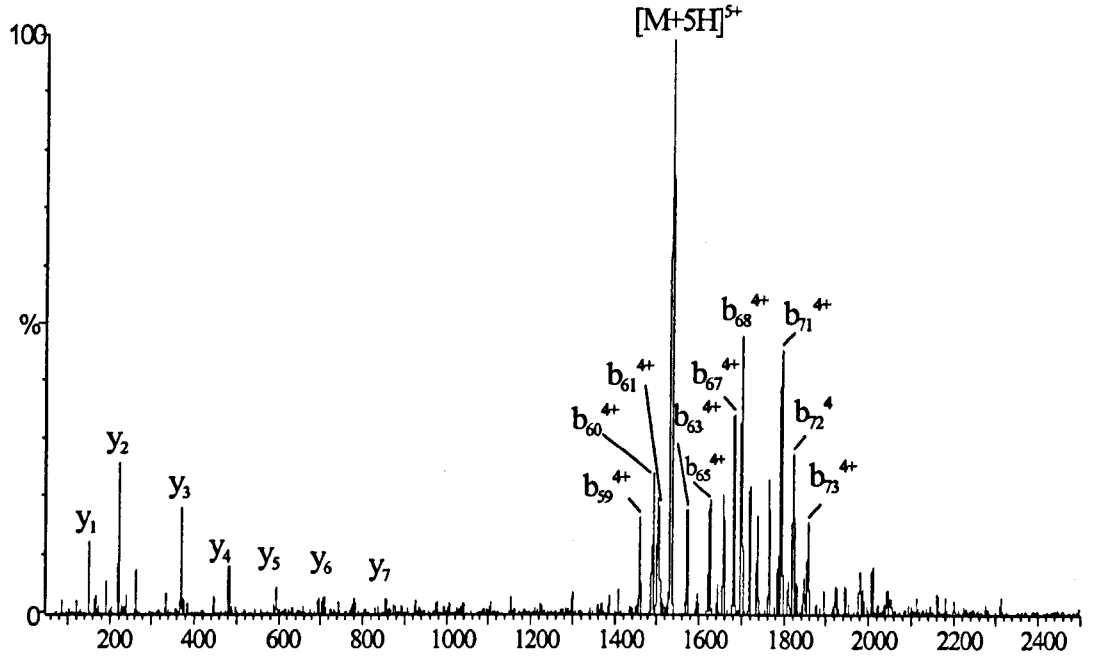


Figure 4.4. MS/MS spectrum of the $[M+5H]^{5+}$ ion of the ceroid lipofuscinosis protein.

This suggests that the 7666.2 Da species present in the ESI-MS spectrum assigned as being oxidation arises from the C-terminal methionine amino acid residue of the CLP The most prominent peaks in the MS/MS spectrum correspond to a series of fragment ions from b_{59}^{4+} to b_{73}^{4+}. The spectrum also contains a further three series of peaks of much lower intensity which correspond to b_2^{+} to b_{14}^{+}, b_9^{2+} to b_{29}^{2+} and b_{46}^{3+} to b_{63}^{3+} fragment ions. The sequence coverage obtained by the MS/MS analysis of the $[M+5H]^{5+}$ ion of the CLP is illustrated in Figure 4.5. Overall the MS/MS experiment confirmed approximately 80% of the primary sequence of the CLP.

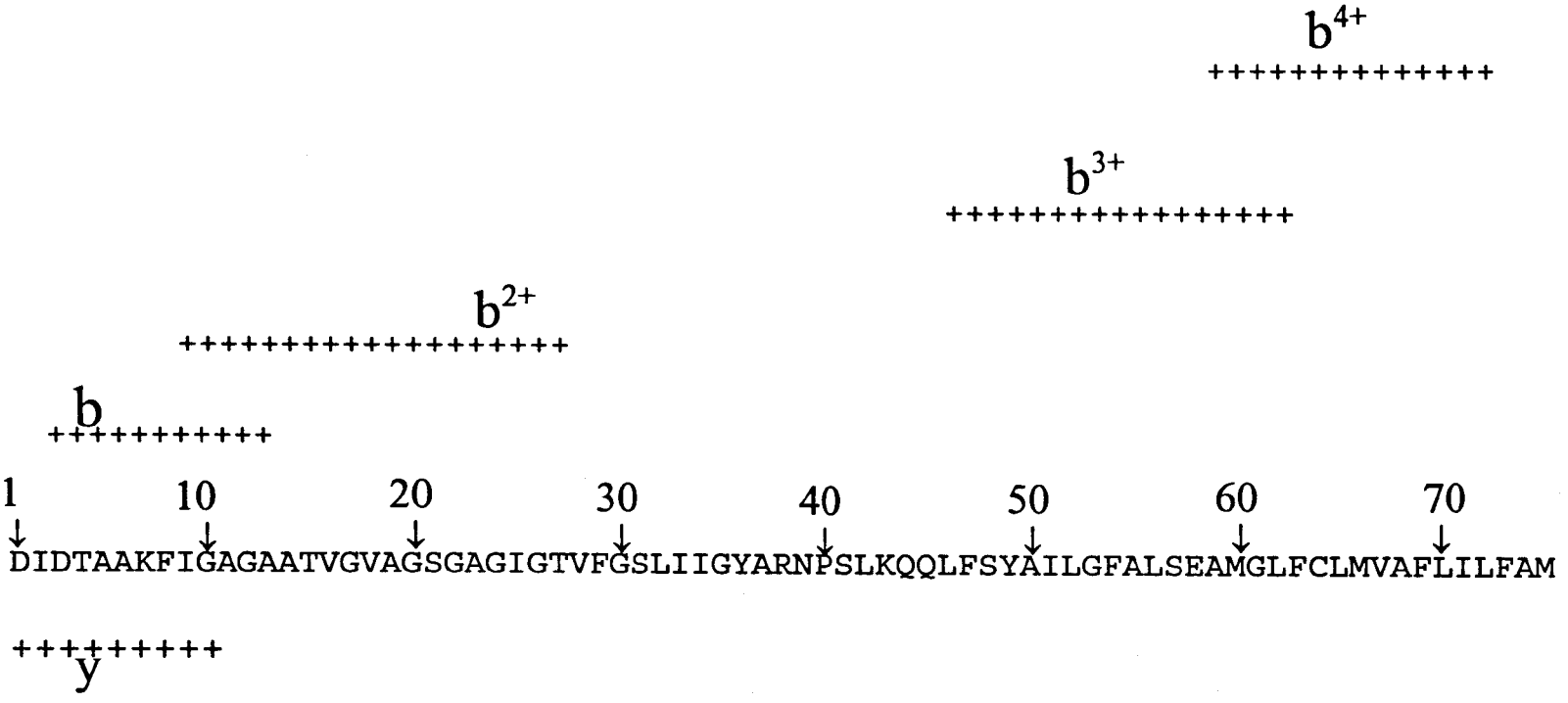


Figure 4.5. Schematic illustrating the fragment ions produced by the CID of the $[M+5H]^{5+}$ ion of the ceroid lipofuscinosis protein.

The partial sequence coverage is consistent with the +42.3 Da modification being present on amino acid residues 30-46. This section of the amino acid sequence is not covered by the MS/MS data. The data are consistent however with acetylation at Lys43, Ser41 or Arg38 or alternatively trimethylation at Lys43. Previous ESI-MS studies [20] have shown through the application of amino group reagents such as 1-fluoro-2,4-dinitrobenzene and phenyl-isocyanate that Lys-43 is already modified and this amino acid remains the prime candidate for further investigation.

4.3 (b) Bovine heart mitochondrial subunit c.

The next stage in the investigation was to establish whether the +42.3 Da modification present on the CLP was due to Batten disease or was it present upon the undiseased subunit c protein. Bovine heart mitochondrial subunit c was therefore extracted and subjected to ESI-MS analysis as shown in Figure 4.6.

The ESI spectrum of bovine heart mitochondrial subunit c is dominated by a number of peaks corresponding to singly charged ions in the range 600-1100 m/z. These are assigned as being lipids and phospholipids which were also extracted during the purification of the subunit c and are characteristic of ESI spectra of bovine heart mitochondrial subunit c extracted in this manner.

The spectrum shows peaks which correspond to the 6+ to 3+ ions of a protein of RMM 7650.3 Da. This protein is assigned as bovine heart mitochondrial subunit c, confirming that the +42.3 Da is found on subunit c extracted from a healthy animal as well as on the ceroid lipofuscinosis protein (ovine and bovine subunit c share an identical amino acid sequence).

The finding that the +42.3 Da modification, whether it be acetylation or trimethylation or an amino acid residue substitution, is present on both the normal and diseased subunit c is in contrast to the results of Katz *et al* [8]. They have suggested that the CLP contains a trimethylation (TML) of one of its two lysine residues but that the normal subunit c in mitochondrial ATP synthase does not contain elevated amounts of TML.

The oxidised component of the subunit c is considerably more significant than was found in the ESI-MS spectrum of the CLP, indeed the peaks corresponding to the oxidised subunit c are the prominent peaks in the 6+ and 5+ charge state clusters. Invariably the subunit c purified by the standard over night extraction was found to contain high levels of the oxidised component. As oxidation of subunit c may modify the enzymatic function of subunit c in the ATP synthase complex, or its status upon extraction, it is of interest to establish the normal condition of subunit c in mitochondria. A rapid micro extraction procedure was therefore devised where the extraction time was reduced to 20 mins as opposed to the standard over-night extraction. The ESI spectrum of mitochondrial subunit c extracted by the rapid micro extraction procedure is shown in Figure 4.7.

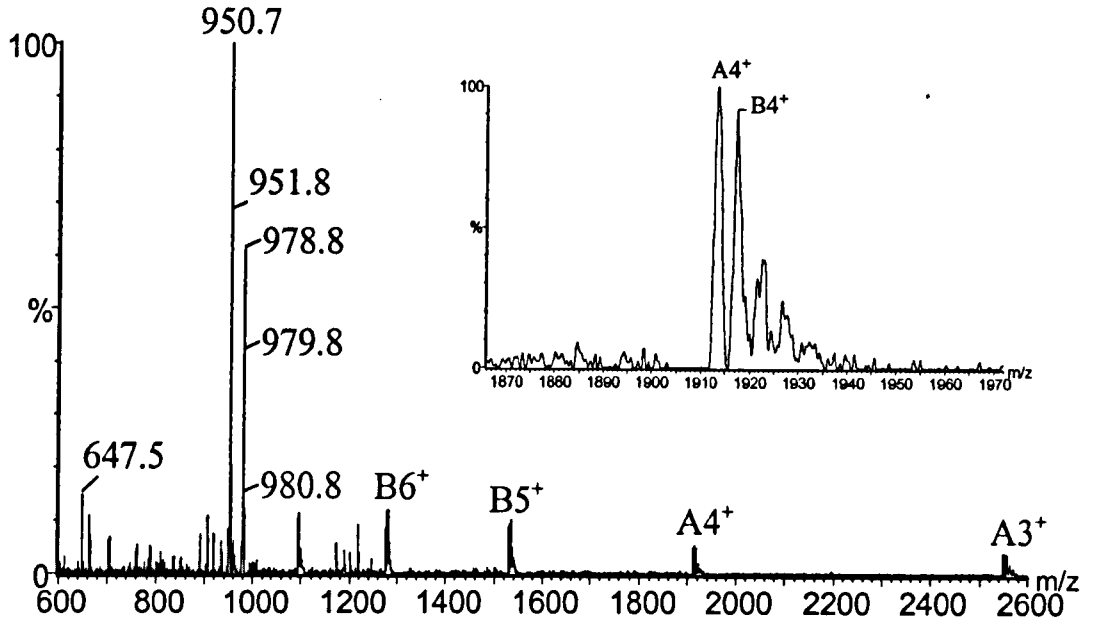


Figure 4.6. ESI spectrum of bovine heart mitochondrial subunit c extracted by the overnight method. Inset shows expanded m/z range which contains the 4+ charge state cluster of subunit c and its oxidised component.

The spectrum in Figure 4.7. shows the characteristic lipids and phospholipids as observed in the ESI spectrum of the mitochondrial subunit c extracted by the standard overnight extraction procedure. Also present are peaks corresponding to the 3+ to 6+ ions of subunit c of measured RMM 7650.2 Da. Examination of the charge state clusters revealed no or very low intensity peaks which correspond to what is believed to be the oxidised form of subunit c. It is apparent from these data that a significant amount

of oxidation of subunit c occurs during the overnight extraction procedure. The measured RMM of the subunit c remains 7650.2 Da suggesting that it contains the same post translational modification or amino acid substitution as the ceroid lipofuscinosis protein and overnight extracted bovine heart mitochondrial subunit c.

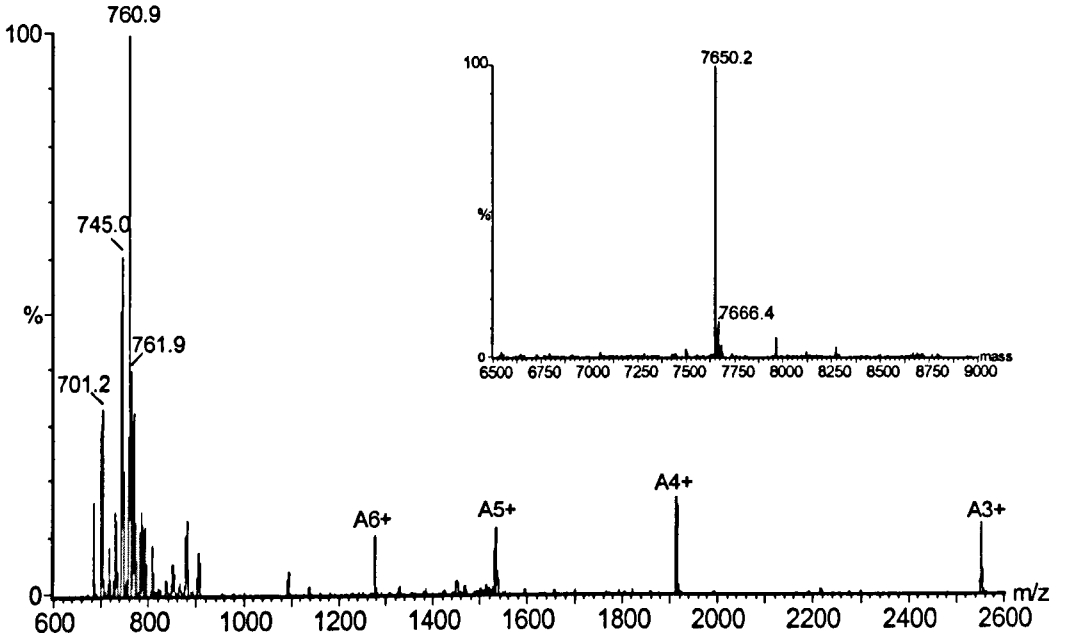


Figure 4.7. ESI spectrum of mitochondrial subunit c extracted by the rapid extraction method. Inset shows the deconvoluted mass spectrum.

4.4 Summary

The work presented in this chapter so far has shown that:

- The hydrophobic ceroid lipofuscinosis protein can be analysed by ESI-MS when dissolved in a solvent system composed of chloroform/methanol/water in the ratio 4/4/1 plus 0.1% HCOOH.
- The CLP has a measured RMM of 7650.3 Da ±0.4 Da
- The RMM calculated from the DNA sequence for subunit c is 7608.0 Da. The difference between the measured and the calculated RMM is +42.3 Da.
- The discrepancy in RMM could be explained by the presence of one acetyl (+42 Da), or a trimethyl (+43 Da) modification of an amino acid residue. Alternatively an amino acid residue substitution such as Gly to a Val (42 Da), Leu to a Arg (42 Da) or Ile to a Arg (43 Da) could account for the discrepancy.
- MS/MS analysis confirmed approximately 80% of the primary sequence of the CLP and was consisted with the modification being present on amino acid residues 30-46.
- Bovine heart mitochondrial subunit c was successfully extracted and analysed by ESI-MS.
- The measured RMM of the bovine heart mitochondrial subunit c was 7650.2 Da, indicating that the same modification is present as is found on the CLP.
- A short term extraction (20 mins) produced subunit c which contained little or no oxidation.

4.5. Interactions with DCCD

From previous investigations [12] the expected CLP-DCCD complexes at Glu-58 are the O-acylisourea (+206 Da) and/or the stable N-acylurea (+205 Da). Interaction of CLP with DCCD in aprotic solvents revealed no significant interaction on ESI/MS analysis and no formation of a CLP-DCCD complex (+205 Da or +206 Da) was detected. Inclusion of water in the incubation medium as described by Fillingame *et al.* [20] was required for the formation of CLP-DCCD complexes and Figure 4.8. shows that other addition products are formed.

Figure 4.8. (A,B,C) shows the ESI/MS spectra obtained after incubation of CLP with DCCD. In addition to the major peak of CLP at 7650 Da and its oxidation product at 7666 Da (+16 Da), the following new species are identified:

(a) The DCCD addition complex (+206 Da) at 7856 Da and its oxidation product at 7872 Da (+206 Da + 16 Da). This is the product expected, presumably by addition to Glu-58, but its levels are variable and it is not a major product.

(b) A major addition product at 7692 Da (+42 Da) and its oxidation product at 7708 Da (+42Da + 16 Da). This is due to a DCCD-catalysed acetylation of CLP utilising the acetate in the ammonium acetate buffer (see below).

(c) A major addition product at 7898 Da (+206 Da + 42 Da) and its oxidation product at 7914 (+206 Da + 42 Da + 16 Da) which is the result of DCCD addition (+206 Da) to the initial acetylated product at 7692 Da (+42 Da).

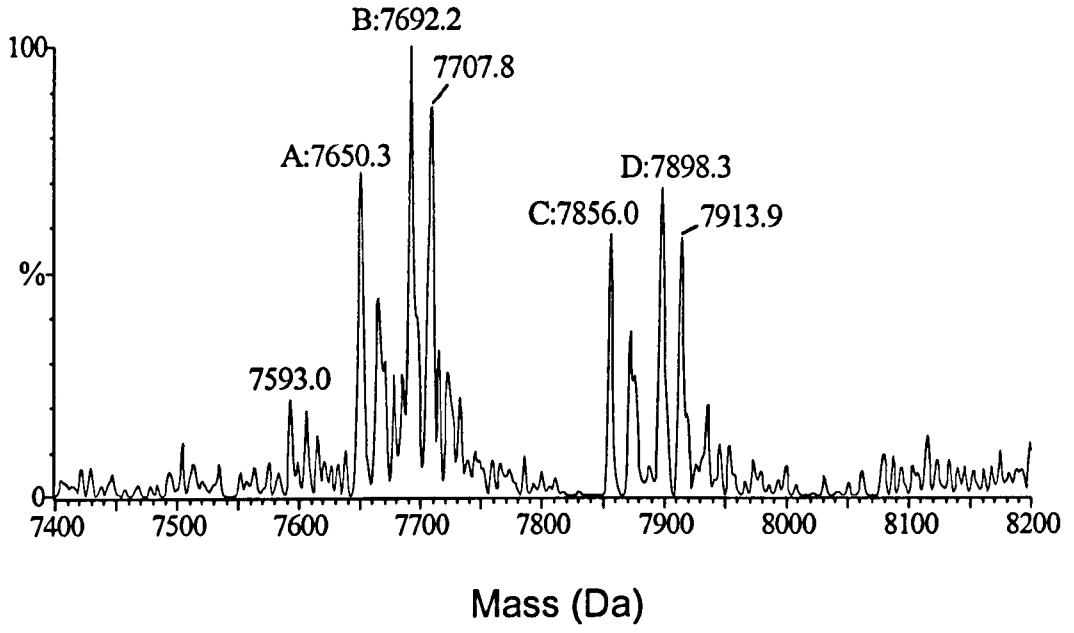


Figure 4.8.A. Deconvoluted mass spectrum of CLP incubated with DCCD at a mole ratio of 10:1 for 18 hours at 4°c. CLP dissolved in C/M/W containing no TFA.

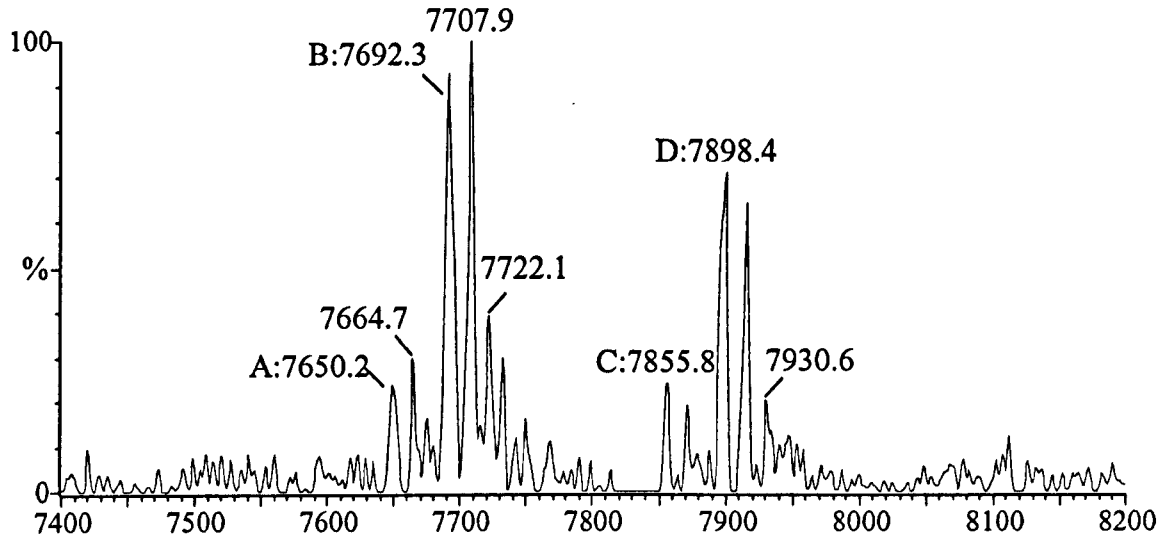


Figure 4.8.B . Deconvoluted mass spectrum of CLP incubated with DCCD at a mole ratio of 10:1 for 18 hours at 4°c. CLP dissolved in C/M/W containing 0.1 %TFA.

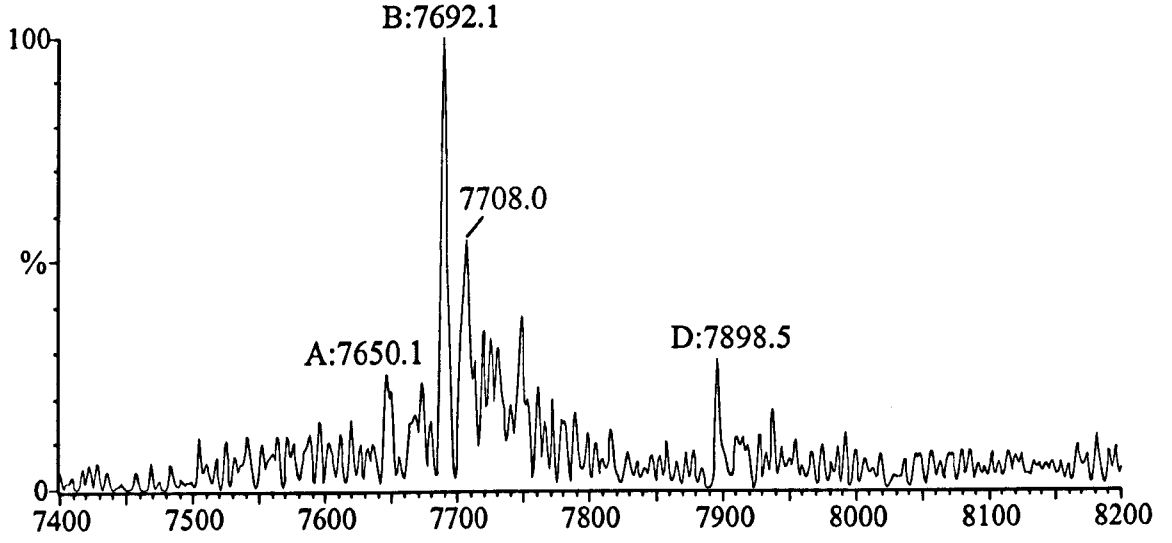


Figure 4.8.C . Deconvoluted mass spectrum of CLP incubated with DCCD at a mole ratio of 10:1 for 18 hours at 4°c. CLP dissolved in C/M/W containing 1.0% TFA.

Kinetic studies of the interaction of DCCD with CLP under these conditions (high acetate concentration, high DCCD:CLP ratio of 100:1) indicate that the primary reaction is the DCCD-catalysed acetylation leading to the formation of the species at 7692 Da (+42 Da) and its oxidation product. This is followed by DCCD addition (+206 Da) to give the acetylated-CLP-DCCD product and, exceptionally, a further DCCD addition (+206 Da). The initial, apparently direct, formation of CLP-DCCD (7856 Da) is a relatively minor reaction pathway but becomes more significant at very high DCCD:CLP ratios. In view of the predominance of the DCCD-catalysed acetylation reaction, however, it is difficult to determine whether the observed CLP-DCCD complex formation occurs directly or whether it is a deacylation or transacylation product of acetyl-CLP-DCCD (7898 Da).

Effects of pH and DCCD:CLP ratios on DCCD interaction.

Increased acidity of the incubation medium leads to a decline in the level of the CLP-DCCD complex at 7856 Da such that at 1% TFA (Figure 4.8C), the acetylated-CLP (+42 Da) is the predominant product with only minor amounts of the CLP-DCCD-acetylated complex (+206 Da +42 Da). Little or no DCCD-CLP complex (+206 Da) is present. The major product at low DCCD:CLP ratios (10:1 or less) is the acetylation product at 7692 Da (+42 Da) and its oxidation products. Formation of less than 5% of the DCCD-CLP complex (+206 Da) or the acetylated-CLP-DCCD complex (+42 Da + 206 Da) is observed, even on complete conversion of CLP. Conversely, at very high DCCD:CLP ratios (>200:1), complete modification of CLP has been observed leading

to two major products: invariably one at 7898 Da (+42 Da +206 Da) and occasionally a novel species at 8104 Da (+42 Da +206 Da +206 Da) with their usual oxidation products (data not shown). Under these conditions the yield of initial acetylation product (+42 Da) is markedly reduced. The second DCCD product has not previously been described and its interaction site has not been established.

In all studies involving the formation of CLP-DCCD adducts, the mass increase is +206 Da in the large majority of cases. Although this suggests that the product is the O-acylisourea derivative, (+206 Da), the resolution achievable with the "Quattro II" mass spectrometer is insufficient to allow a definitive demonstration of the formation, if any, of the N-acylurea derivative (+205 Da).

DCCD-catalysed Acetylation and the Acetylation Site

To establish that the source of the acetyl group (+42Da) is the ammonium acetate buffer, control experiments were carried out with CLP that had been extensively washed by the method of Folch *et al.* [21] to reduce the acetate content. Such preparations show little or no formation of the 7692 Da species (+42 Da) on incubation with DCCD but this reaction is restored on addition of ammonium acetate. Similarly, ammonium acetate restores the DCCD-catalysed acetylation reaction to CLP prepared in the absence of ammonium acetate, thus confirming that the ammonium acetate buffer is the source of the acetyl group (+42 Da).

The possible role of Cys-64 as an acetylation site on CLP or as a transient site in a transacetylation reaction has been investigated by titration of Cys-64 with 2,3-dimethoxy-6-methyl,1,4-benzoquinone (UQ_0) to form a CLP-UQ_0 complex (+180 Da) as previously described [22]. Incubation of CLP-UQ_0 with DCCD leads to the formation of the acetylated product (+180 Da + 42 Da) and the acetylated-CLP-UQ_0-DCCD product (+180Da + 42 Da +206 Da) as shown in Figure 4.9. It is concluded that Cys-64 is unlikely to be involved in the DCCD-catalysed acetylation reaction or any consequent transacylation reaction.

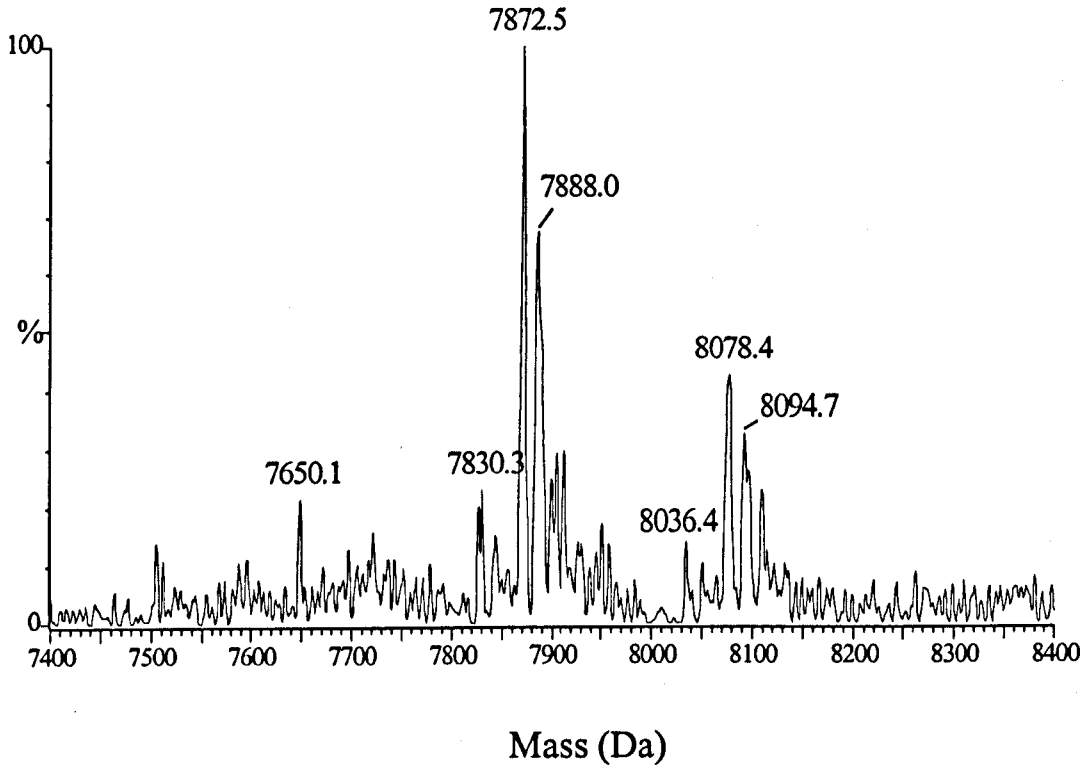


Figure 4.9. Deconvoluted mass spectrum showing the effects of 2,3-dimethoxy-6-methyl-1,4-benzoquinone on the reaction of CLP with DCCD.

CID analysis of the acetylated CLP species of RMM 7692 Da was performed on the $[M+5H]^{5+}$ ion of the modified species. This analysis establishes that the acetylation is associated with one of the first nine amino acids in CLP but it was not possible to define the particular amino acid involved. Lys-7 and N-terminal aspartate appear to be the prime candidates.

Other DCCD-catalysed acylations

The DCCD-catalysed acylation is not limited to the acetylation reaction (+ 42 Da). Under appropriate conditions, DCCD-catalysed acylation of CLP by the fatty acids myristic, palmitic, oleic and stearic acids has been observed. Incubation of CLP with DCCD and the ammonium salts of fatty acids under standard conditions (CLP diluted 1:1 with CMW; 50 mM ammonium acetate) led to the formation of acetyl-CLP (+42 Da) only. No fatty acid or DCCD adducts were observed.

Water washing of CLP as described above [21] reduces the ammonium acetate content to minor levels. Incubation of water washed CLP with ammonium salts of fatty acids (myristate, palmitate, oleate and stearate) led to the formation of the expected fatty acid adducts as well as the CLP-DCCD adduct (+206 Da). Acylation of CLP by fatty acids was markedly slower than acetylation with acetate.

4.6. Conclusions

On interaction of CLP with DCCD, the expected product CLP-DCCD (7865 Da) is formed, presumably by formation of the O-acylisourea product at Glu-58 [12]. It is not the major product, however; these are acetyl-CLP (7692 Da) and acetyl-CLP-DCCD complex (7898 Da). In addition, evidence for a second DCCD adduct is presented, i.e. CLP-$(DCCD)_2$ (+42Da +206Da +206Da) and the acetylated adduct acetyl-CLP-$(DCCD)_2$ (+42Da +206Da +206 Da). The location of the second DCCD has not been established.

A preliminary kinetic analysis indicates that the primary reaction product is acetyl-CLP (+42 Da) and not CLP-DCCD (+ 206 Da). Titration of Cys-64 with UQ_0 (Figure.4.9) has no effect on the DCCD interaction, but elimination of Cys-64 as a site of acetylation and/or transacylation will require evidence for complete conversion of CLP to CLP-Q_0 before addition of DCCD. CID analysis of acetyl-CLP indicates that the probable site of acetylation is Lys-7 or N-terminal aspartate, an acylation site which in F_0 is located on the M-side of the inner membrane.

The demonstration that the predominant initial reaction is acetylation raises the question as to whether prior acylation of CLP is required for DCCD addition as opposed to the direct addition of DCCD to Glu-58 of CLP. Acylation of CLP may result in a conformational change which facilitates the formation of the O-acylisourea product at

Glu-58, thus providing an explanation for the function of "excess" DCCD [12]. The product, acetyl-CLP-DCCD, would then be deacylated to give CLP-DCCD. There is no experimental evidence available to support this view, however, and it is not possible at present to distinguish between these two possibilities.

The novel observations described above are model reactions with CLP (subunit c) whose relevance to subunit c function in intact membrane-bound ATP synthase (F_1F_0 ATP synthase) remains to be established. Nevertheless, they establish possible reactions and experimental systems using F_0 inhibitors which were not readily available for investigation by use of previous analytical methods. Studies of membrane-bound ATP synthase are now feasible as subunit c and any derivative forms can be extracted by chloroform/methanol solvent mixtures and then analysed by means of ESI/MS. Of immediate interest is the question of whether the subunit c-DCCD complex isolated from mitochondria is or can be acylated in an analogous fashion to acetyl-CLP-DCCD.

Of more general interest is the question of whether acylated forms of subunit c are normal mitochondrial constituents and whether there is an enzymic acylating system which operates on the M-side of the mitochondrial inner membrane analogous to the DCCD-catalysed acylation. It has already been established that normal mitochondrial subunit c carries an unidentified 42 Da post-translational modification associated with Lys-43 [21], a possible acylation site which in F_0 is located on the C-side of the inner membrane.

4.7 References

(1) W. Zeman, P. R. Dyken, *Pediatrics,* **44**, 570 (1969).

(2) P. R. Dyken, *Am. J. Med. Genet. Suppl.*, **5**, 69 (1988).

(3) J. A. Rider, D. L. Rider, *Am. J. Med. Genet. Suppl.*, **5**, 21 (1988).

(4) D. N. Palmer, G. Barns, D. R. Husbands, R. D. Jolly, *J. Biol. Chem.*, **261**, 1773 (1986).

(5) I. M. Fearnley, J. E. Walker, R. D. Martinus, R. D. Jolly, K. B. Kirkland, G. J. Shaw, D. N. Palmer, *Biochem. J.*, **268**, 751 (1990).

(6) D. N. Palmer, I. M. Fearnley, J. E. Walker, N. A. Hall, B. D. Lake, L. S. Wolfe, M. Haltia, R. D. Martinus, R. D. Jolly., *Am. J. Med. Genet. Supp.*, **42**, 561 (1992).

(7) R. D. Jolly, R. D. Martinus, D. N. Palmer, *Am. J. Med. Genet. Supp.*, **42**, 609 (1992).

(8) M. L. Katz, M. Rodrigues, *Am. J. Pathol.*, **138**, 323 (1991).

(9) D. N. Palmer, S. L. Bayliss, P. A. Clifton, V. J. Grant, *J. Inher. Meatab. Dis.*, **16**, 292 (1993).

(10) D. D. Tyler, *Membr. Struct. Funct.*, **5**, 117 (1984).

(11) W. Sebald, W. Machleidt, E. Wacheter, *Proc. Natl. Acad. Sci. USA*, **77**, 785 (1980).

(12) W. Sebald and J. Hope, In 'Current Topics in Bioenergetics', D. R. Sandi (Ed.). Academic Press, New York, Vol. 12, (1984) pp.1.

(13) R. H. Fillingame, In '*Current Topics in Bioenergetics*', D. R. Sandi (Ed.), Academic Press, New York, Vol. 11, (1981). pp. 35

(14) R. H. Fillingame, *Biochim. Biophy. Acta*, **1101**, 240 (1992).

(15) E. M. Ryan, A. Buzy, D. E. Griffiths, K. R. Jennings, D. N. Palmer, *Biochem. Soc. Trans.*, **24**, 289S, (1996).

(16) P. A. Schindler, A. Van Dorsselaer, A. M. Falick, *Anal. Biochem.*, **213**, 256, (1993).

(17) D. N. Palmer, R. D. Martinus, R. D. Barns, R. D. Reeves, R. D. Jolly, *Am. J. Med. Suppl.*, **5**, 141, (1988).

(18) S. M. Medd, J. E. Walker, R. D. Jolly, *Biochem. J.*, **293**, 65, (1993).

(19) A. Buzy and D. E. Griffiths, unpublished results

(20) R. H. Fillingame, M. E. Girvin, Y. Zhang, *Bio-chem. Soc. Trans.*, **23**, 760 (1995).

(21) J. Folch, M. Lees, G. H. Sloane Stanley, *J. Biol. Chem.*, **226**, 497 (1957)

(22) A. Buzy, E. M. Ryan, K. R. Jenninngs, D. N. Palmer, D. E. Griffiths, *Rapid Commun. Mass Spectrom.*, **10**, 790, (1996).

Chapter 5

Characterisation Of The Soluble Methane Monooxygenase Hydroxylase β-Subunit

5.1. Introduction

Methane monooxygenases (MMOs) are multicomponent enzyme systems present in methanotrophic bacteria [1]. These bacteria play an essential role in cycling carbon in the biosphere by consuming methane produced in anaerobic sediments [2]. The first step in this process is catalysed by MMO which catalyses the conversion methane to methanol [1].

$$CH_4 + O_2 + NADH + H^+ \xrightarrow{MMO} CH_3OH + H_2O + NAD^+$$

In addition, MMO has a very wide substrate specificity, that is to say it also catalyses the oxygenation of a range of organic compounds including alkanes, alkenes, ethers, alicyclic, aromatic and heterocyclic compounds [1, 3-5]. This broad substrate specificity justifies the study of the enzyme because of its possible use for the biotreatment of hydrocarbon pollutants [6]. Detailed structural investigations of MMO have, therefore, been undertaken in order to understand the mechanism of its enzymatic activity. Furthermore, an understanding of the biological mechanism may lead to the design of direct catalysts that could replace the energetically demanding chemical route to methanol that requires the use of three independent catalysts.

A number of methanotrophs can express a membrane bound particulate (particulate MMO) or a soluble (soluble MMO) form of the enzyme.

The particulate enzyme is less well understood although lately progress has recently been made in its purification and characterisation [7,8]. The soluble enzymes from *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b have been most studied and reviews of both have appeared recently [9,10]. The soluble MMOs from these organisms are similar three component systems made up of an hydroxylase, a reductase and a regulatory protein, B. Electrons from NADH are transferred to the hydroxylase *via* the Fe_2S_2 and FAD cofactors of the reductase. The hydroxylase has two copies of three subunits in an $\alpha_2\beta_2\gamma_2$ configuration. The Fe-(OH)-Fe site, at which both O_2 and CH_4 are activated, resides in the a-subunit. On binding to the hydroxylase, protein B is believed to cause a conformational change in the former which somehow regulates electron flow to the diiron site [9].

A 2.2 Å resolution X-ray crystallographic investigation of the soluble MMO hydroxylase from *M. capsulatus* (Bath) [11] revealed some errors in the published gene sequence [12]. Therefore a study of the sMMO hydroxylase was undertaken by means of electrospray ionisation mass spectrometry. This earlier investigation of the sMMO, performed by A. Buzy [13], revealed that the measured RMMs of the three subunits of the hydroxylase (α,β and γ) disagreed with the predicted RMMs, calculated using the gene sequence as illustrated in Table 5.1.

Protein subunit	Measured molecular mass (Da)	Expected Average molecular mass (Da)	Measured-Expected (Da)
α	60 518.3 ± 5.8	60 501.3	+ 17.0
α'	59 774.9 ± 6.0	59 757.5	+ 17.4
α''	59 973.9 ± 8.4	59 956.7	+ 17.2
β	45 001.7 ± 2.1	44 875.6	+ 126.1
γ	19 715.6 ± 0.9	19 710.5	+ 5.1

Table 5.1. The measured and predicted relative molecular masses of the hydroxylase subunits

The discrepancies in RMM of the α and γ subunits were located and identified using a standard method as shown in Figure 5.1. The first aim of this method is to produce a low number of cleavage peptides one of which, it is hoped, will retain the discrepancy in RMM. For this reason a cleavage reagent is chosen which cleaves only at a limited number of amino acids. Cyanogen bromide (CNBr) is a common choice as it only cleaves at the C-terminal side of methionine. The next step is to isolate the cleavage peptide of interest by HPLC before subjecting it to an enzymatic digestion to produce

much smaller digest peptides (typically of RMM less than 1500 Da), which are amenable to MS/MS analysis.

CNBr cleavage (simple mixture)

↓

Isolation of peptide containing discrepancy in RMM by HPLC

↓

Enzymatic digestion of peptide

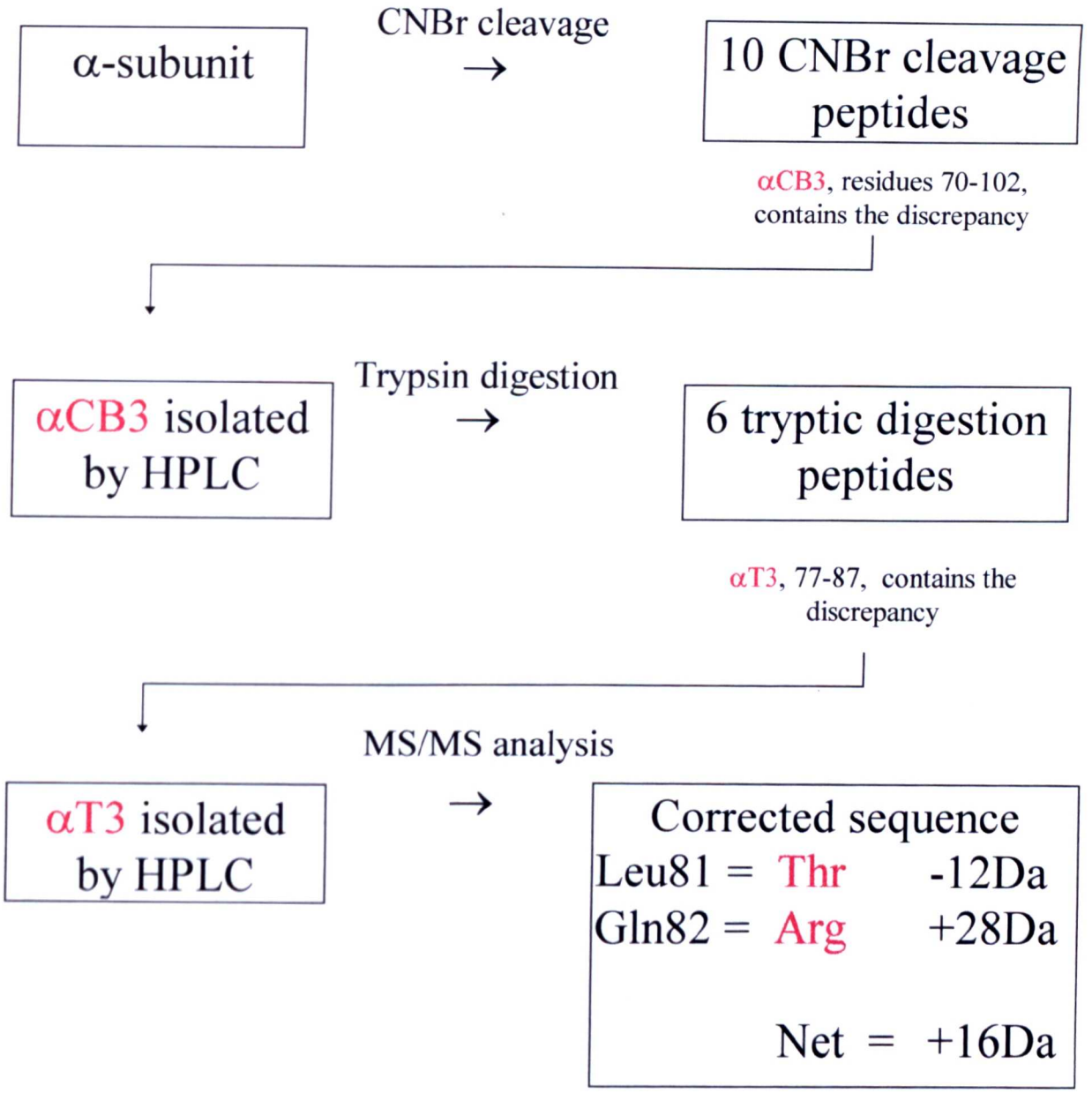
↓

MS/MS analysis

Scheme 5.1. Schematic showing standard method for location and identification of discrepancy in RMM of protein

α-subunit

The discrepancy of +17 Da for the α-subunit was shown to be present in residues 77-87 by means of the standard method to be present in residues 77-87. MS/MS analysis of the $[M+2H]^{2+}$ ion of this peptide indicated that residue Leu81 should be replaced by Thr and Gln82 by Arg (see scheme 5.2).



Scheme 5.2. Schematic and results of standard method of location and correction for the sequence error of the α-subunit.

Furthermore, three proteins of RMM 60518, 59775 and 59974 Da were observed. The first was assigned as the α-subunit the second and third as the α-subunit having been cleaved at residues 8 and 6 respectively.

γ-subunit

The sequence of the γ-subunit was successfully corrected using an method identical to that used for the α-subunit. The corrections were as follows;

Val4	→	Ile/Leu	=	+14 Da
His3	→	Gln	=	-9 Da
		Net	=	+5 Da

Figure 5.1. Sequence corrections made for the hydroxylase γ-subunit.

β-subunit

Using the standard method the discrepancy of +126.1 Da present in the β-subunit was located in a tryptic peptide consisting of residues 141-146. MS/MS analysis of the $[M+H]^+$ ion of the 141-146 peptide, performed on the "Quattro II" instrument, indicated a sequence of the type DEASCNR (predicted sequence TSSCNR) which would have fully accounted for the anomaly in RMM. Subsequent studies revealed, however, that

the CID spectrum contained ions of the products of two digest peptides of closely similar mass, that consisting of residues 141-146, m/z 792.4, and that of residues 91-96, m/z 794.4.

The transmission and subsequent CID of the two peptides was due to the resolution of Q1 being lowered to compensate for the low sensitivity of the experiment achieved by using the standard method. A revised method was therefore devised which would increase the sensitivity of the experiment, this consisted of omitting the first two step of the standard method (Scheme 5.3.). It was recognised that a tryptic digestion of the β-subunit would produce a complex digestion mixture (46 predicted digest peptides). The prior data concerning the location of the mass discrepancy would however greatly simplify the identification of the relevant peptide in the complex digest mixture.

Tryptic digest of intact β-subunit

(Complex mixture)

↓

MS/MS analysis

Scheme 5.3. Revised method for location and identification of discrepancy in RMM of the hydroxylase β-subunit.

5.2. Experimental

Purification of the hydroxylase

Growth of the *M.capsulatus* (Bath) organism was as previously described [14]. The following modifications of the published procedures for the purification of the hydroxylase were employed. After the initial ion exchange step, the hydroxylase was applied to a Superdex 200 gel filtration column. Final purification was *via* elution of pure hydroxylase from a Mono Q column using a 0-30% gradient of 1M NaCl. Prior to injection into the mass spectrometer, buffer was removed from all protein solutions on a fast desalt column.

Hydroxylase subunit separation

The subunits of the hydroxylase were separated by fast-protein liquid chromatography (FP-LC) using a Pro-RPC 15μ column. The gradient ran from H_2O to 90% CH_3CH containing 0.1% TFA. The β-subunit was eluted in approximately 60% CH_3CN containing 0.1% TFA.

Mass Spectrometry

The majority of mass spectra were recorded by means of a "Quattro II" QhQ tandem mass spectrometer (Micromass UK Ltd, Altrincham, UK) equipped with an atmospheric pressure ionisation source operated in the nebulizer-assisted electrospray mode. The potential of the electrospray needle was set at 3.5 kV and the extraction cone voltage

was varied linearly from 30V at m/z 600 to 90V at m/z 2600. The source temperature was held at 60°C. The carrier solvent used was a 1:1 mixture of CH_3CN/H_2O containing 0.1% HCOOH. The samples were infused using a syringe pump (Harvard model 11, South Natic, USA) at a flow rate of 5μL/min. Calibration was carried out using a solution of horse heart myoglobin or sodium iodide. Mass spectra were acquired over the range m/z 600-2000 (m/z 200-2000 for digest spectra) during a 10s scan and by operating the data system in the multichannel acquisition (MCA) mode.

For low energy MS/MS experiments, precursor ions were selected by Q_1 and a potential difference of 30 V between Q_1 and the hexapole collision cell produced collision energies of 30z eV where z is the number of charges on the precursor ion. The indicated argon collision gas pressure was 3.2×10^{-3} mbar and Q_2 was scanned over the range m/z 50-900 to produce the product ion spectrum. The Q_2 was calibrated using a mixture of NaI (2μg/μl) dissolved in a solution of water/propan-2-ol (1:1). All data were processed by means of the MassLynx data system. The reconstructed mass spectra were obtained by using the "Max Entropy" algorithm.

Several spectra were acquired on the "Q-ToF" mass spectrometer (Micromass UK, Altrincham, Cheshire) in which ions produced in an ESI source are transported (in MS mode) by quadrupole and hexapole ion guides to an orthogonal time of flight (TOF) mass analyser. In MS/MS mode the quadrupole mass analyser is used to select precursor ions which are fragmented in the hexapole collision cell to produce fragment

ions which are analysed by the TOF mass analyser. This instrument has a sensitivity two orders of magnitude greater than that of the "Quattro II" at a mass resolving power of approximately 5000, thereby increasing the reliability of sequencing a small number of tryptic peptides.

Tryptic digestion

Approximately 1nmol of the β-subunit in 60% CH_3CN containing 0.1% TFA was taken to near dryness by ultrafiltration in a micron 10 kDa microconcentrator (Amicon Inc., Beverley, MA, USA). The protein was reconstituted in 100 μL of H_2O to which was added 10 μL of CH_3CN, 10 μL of 1M NH_4HCO_3 and 1/50 (w/w trypsin/protein) of trypsin. The digestion was allowed to proceed at 37°C for 8 h. The digestion mixture was diluted to approximately 2 pm/μL in CH_3CN/H_2O (1:1) plus 0.1% HCOOH before analysis by ESI/MS and MS/MS.

Reduction and carboxyamidomethylation of the intact β-subunit

10 μL of 1 M NH_4HCO_3, 5 μL of 100 mM DTT and 5 μL of 200 mM iodoacetamide were added to approximately 1 nmol of β-subunit (dissolved in 60% CH_3CN plus 0.1% TFA) and incubated at 37°C for 30 min. The mixture was then taken to near dryness by ultrafiltration in a micron 10 kDa microconcentrator and washed twice with 100 μL of 0.1% HCOOH. The reduction and carboxyamidomethylation of the protein was confirmed by ESI/MS with the remainder of the peptide being subjected to a tryptic digestion as described above.

5.3 Results and discussion

5.3 (a) Hydroxylase

Figure 5.2. shows the ESI spectrum of the sMMO hydroxylase. Under the low pH conditions of the spraying solvent the non-covalently bound subunits of the hydroxylase show individual charge state distributions due to each subunit. Two charge state distributions of multiply charged ions which corresponds to proteins of RMM 19 714.9 ± 1.2 and 45 001.4 ± 2.0 are clearly visible.

These proteins are attributed to the γ and β subunits respectively. A third charge state distribution is present at higher m/z values at much lower intensity. Maximum entropy deconvolution of the ESI spectrum of the sMMO hydroxylase produces a much simpler spectrum which clearly identifies the three expected subunits. The "MaxEnt" spectrum of the sMMO is shown in Figure 5.3. This spectrum has been processed so that the heights of the peaks are proportional to the concentrations of the components.

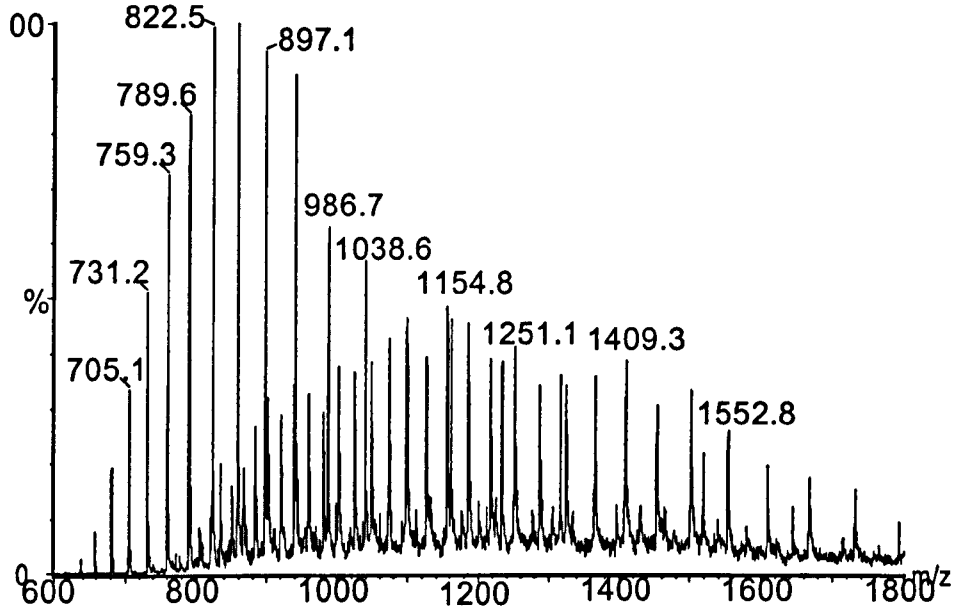


Figure 5.2. ESI spectrum of the sMMO hydroxylase.

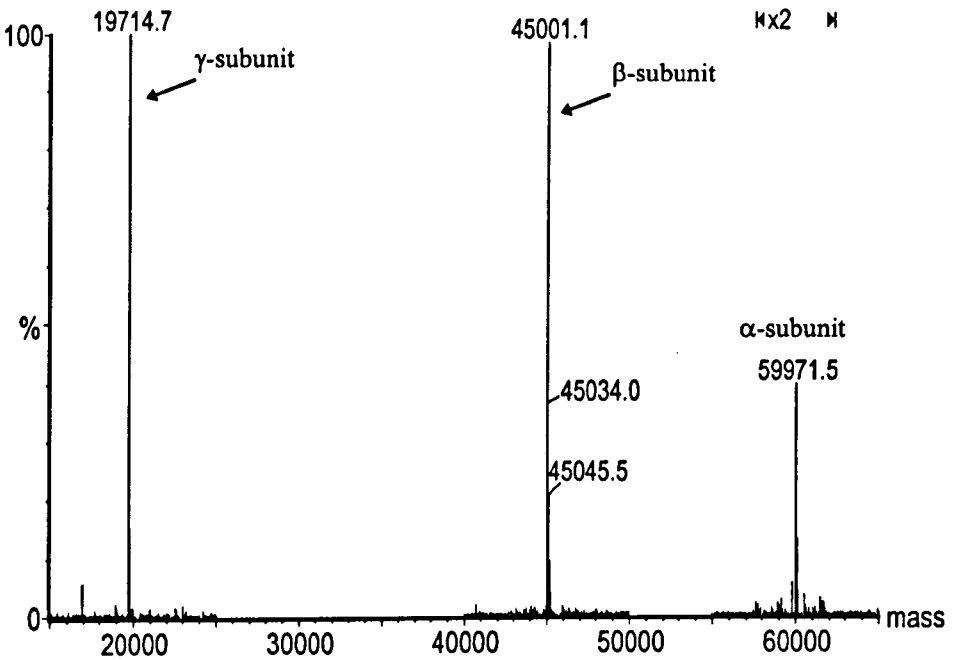


Figure 5.3. "MaxEnt" deconvoluted mass spectrum of the sMMO hydroxylase

The α-subunit is of much lower intensity than the other two which is unexpected as the subunit are known to be in a 1:1:1 ratio.

5.3 (b) β-subunit

In order to study the β-subunit the components of the hydroxylase were separated by fast protein liquid chromatography. The ESI spectrum of the β-subunit is shown in Figure 5.4.

The spectrum shows a single charge state distribution with charges ranging from +23 to +65. The "MaxEnt" spectrum, inset, indicates that the measured RMM of the β-subunit to be 45001.9 Da ±1.9 Da.

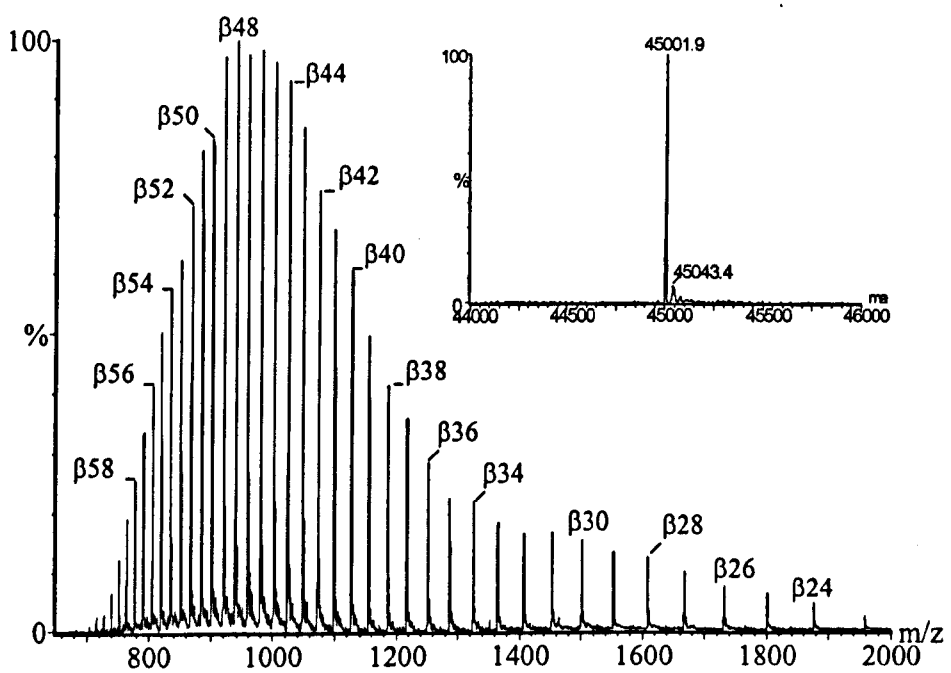


Figure 5.4. ESI spectrum of the β-subunits, inset shows the "MaxEnt" deconvoluted mass spectrum.

This disagrees with the predicted RMM, calculated from the gene sequence, of 44875.6 Da, by +126.3 Da. In order to locate the discrepancy within the protein, whilst retaining adequate sensitivity, a tryptic digest of the β-subunit was undertaken. Prior knowledge of the primary sequence allows the expected tryptic digest peptides to be calculated in advance. In this instance a complete tryptic digestion of the β-subunit should produce 46 digest peptides.

5.3 (c) Tryptic digestion of the β-subunit

Figure 5.5. shows the ESI spectrum of the tryptic digest of the intact β-subunit.

The notation of the type Tn indicates the tryptic digest peptide. The good agreement between measured and expected RMM values allowed the majority of the peptides to be assigned as β-subunit digest peptides as shown in table 5.2. The tryptic peptides T44 and T41 were not located in the digest mixture, but were present as part of larger digest peptides (T44-45, and T40-41) due to incomplete digestion.

Several other digest peptides comprising of a single arginine residue were undetected. Overall the tryptic digestion of the native β-subunit confirmed 90% of the primary sequence.

Only two peptides, of more than four residues, were undetected in the digestion mixture. Neither peptide T35 nor T21 (RMM 2611.7 and 666.3 Da respectively) were detected at their expected m/z values. A peptide was, however, detected at 792.4 Da, confirming

that the discrepancy, now measured more precisely as 126.1 Da, is located in the T21 peptide (residues 141-146 TSSCNR).

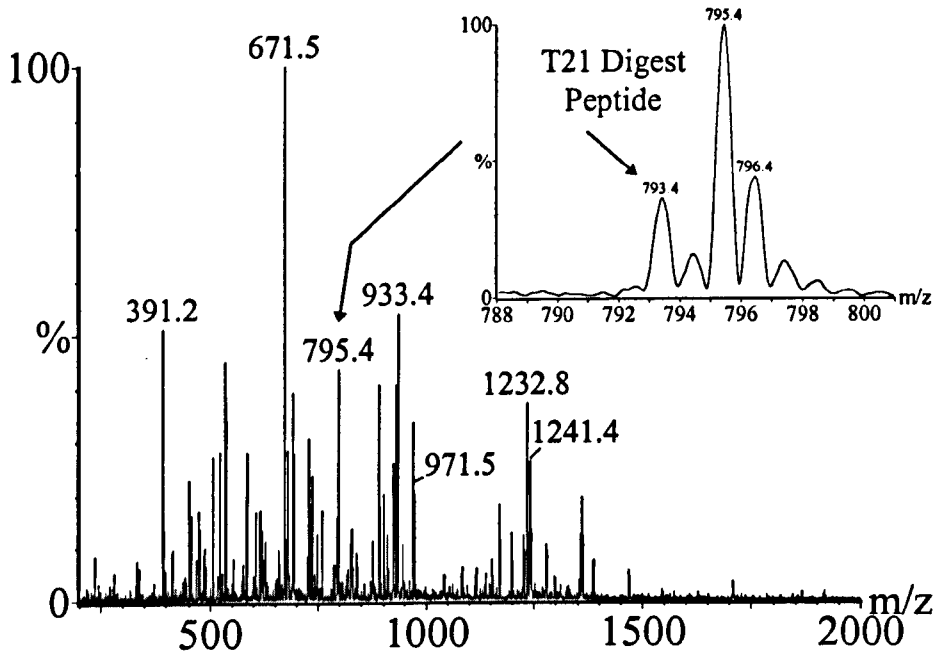


Figure 5.5. ESI spectrum of the tryptic digestion of the β-subunit, inset shows an expanded region of the ESI spectrum which contains the T21 tryptic peptide $[M+H]^+$ ion.

Table 5.2. Tryptic digest map of the β-subunit

Digest Peptide	Residues	Predicted RMM (Da)	Measured RMM (Da)	Difference (Da)
T1	1-6	691.3	691.4	+0.1
T2	7-7	174.1	Not detected	-
T3	8-8	174.1	Not detected	-
T4	9-21	1356.7	1356.8	+0.1
T5	22-33	1237.6	1237.6	0.0
T6	34-41	969.5	969.5	0.0
T7	42-43	332.2	332.2	0.0
T8	44-44	174.1	Not detected	-
T9	45-74	3412.7	3413.1	+0.4
T10	75-90	1842.8	1843.4	+0.5
T11	91-96	794.4	794.4	0.0
T12	97-98	311.2	Not detected	-
T13	99-102	499.2	Not detected	-
T14	103-103	174.1	Not detected	-
T15	104-110	899.5	899.5	0.0
T16	111-112	261.1	Not detected	-
T17	113-117	689.3	689.5	+0.2
T18	118-121	553.3	553.3	0.0
T19	122-133	1353.7	1353.7	0.0

T20	134-140	874.4	874.5	+0.1
T21	**141-146**	**666.3**	**792.4**	**+126.1**
T22	147-167	2448.7	2447.8	0.0
T23	168-175	889.5	889.5	0.0
T24	176-185	1168.6	1168.7	+0.1
T25	186-196	1328.7	1328.9	+0.2
T26	197-201	534.3	534.3	0.0
T27	202-214	1358.7	1358.7	0.0
T28	215-224	1195.5	1195.6	+0.1
T29	225-227	332.2	332.2	+0.0
T30	228-260	3831.2	3831.0	-0.2
T31	261-261	174.1	Not detected	-
T32	262-266	725.4	725.3	-0.1
T33	267-270	455.3	455.3	0.0
T34	271-291	2463.8	2463.6	-0.2
T35	292-313	2611.7	Not detected	-
T36	314-317	505.3	505.2	-0.1
T37	318-322	604.3	604.3	0.0
T38	323-332	1168.7	1168.7	0.0
T39	333-340	927.5	927.5	0.0
T40	341-348	801.4	801.4	0.0
T41	349-357	1080.5	Not detected	-
T42	358-369	1466.7	1466.8	+0.1
T43	370-373	521.3	521.2	-0.1

T44	374-376	360.1	Not detected	-
T45	377-381	601.3	601.3	0.0
T46	382-388	670.4	670.5	+0.1

Table 5.2. Tryptic digest peptide map of the β-subunit continued

5.3 (d) MS/MS analysis of the T21 trytpic digest peptide of the β-subunit

Without further separation this peptide was submitted for MS/MS analysis. Figure 5.6. shows the product ion spectrum of the $[M+H]^+$ ion of the T21 peptide acquired on the "Quattro II" mass spectrometer.

The spectrum shown in Figure 5.6. indicates a radical departure from the predicted product ions of the T21 peptide (predicted sequence TSSCNR). The peak at m/z 175.1 is clearly identified as a y_1'' fragment ion of a tryptic peptide containing arginine at its C-terminus. The peaks at m/z 548.9 and 677.9 could be assigned as y_4'' and y_5'' fragment ions indicating the presence of D and E at the N-terminus (predicted m/z values for the y_5'' and y_6'' ions are 549.3 and 678.4 respectively), further supported by the presence of the complementary b_2 ion at m/z 245.1.

The intense low m/z immonium ions at 120.0 and 86.1 are evidence of F and I/L being present in the peptide. The peak at m/z 391.8 may therefore be identified as the b_3 ion arising from the sequence DEF fragment and by mass balance, Asp (N) is suggested as

the missing residue. A tentative assignment of the sequence of the 792.4 Da peptide is therefore either DEFL/INR or DEFNL/IR. The ultimate sensitivity of the MS/MS experiment on the "Quattro II" instrument was limited by the fact that lowering the resolution too much would permit the transmission of the T11 (residues 91-96) tryptic peptide of RMM 792.4 m/z.

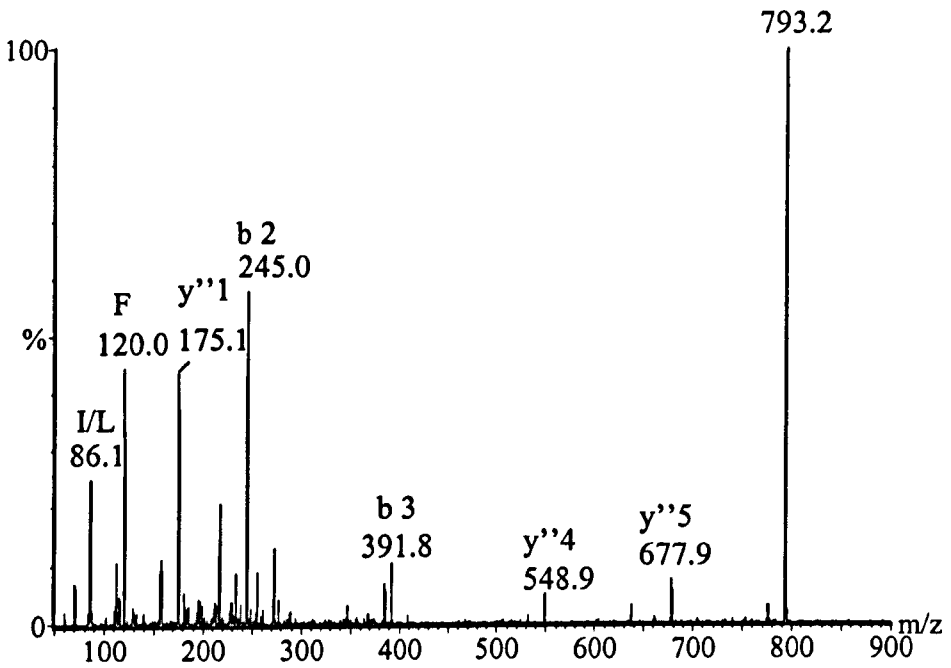


Figure 5.6. MS/MS spectrum of the MH^+ ion of the 792.4 m/z peptide acquired on the "Quattro II" instrument

For this reason the $[M+H]^+$ ion of the 792.4 m/z T21 peptide was also submitted for MS/MS analysis on the "Q-ToF" instrument as it was hoped that the increased

sensitivity and thus mass accuracy would provide a unambiguous assignment of the revised sequence.

Figure 5.7. shows the MS/MS spectrum given by the $[M+H]^+$ ion of the 792.4 Da tryptic peptide T21 acquired on the "Q-ToF" instrument. The improved mass accuracy confirms the assignment made using the "Quattro II" data. The peaks at m/z 505.2 and 402.2 are identified as b_4 and y_2'' ions of the sequence DEFL/IXR. Calculation of mass difference and the presence of a peak at m/z 289.2, attributed to a y_2'' ion, allows the presence of N (Asparagine) to be confirmed (156+114+17+2 Da), giving a final sequence of DEFL/INR. It is not possible to differentiate between leucine or isoleucine in a low energy MS/MS spectrum.

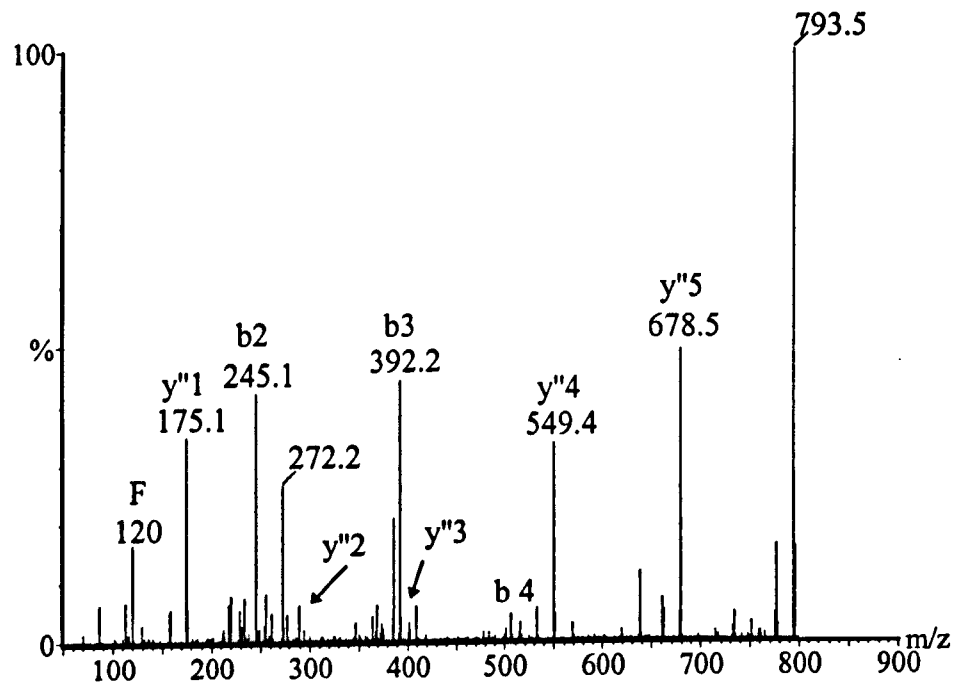


Figure 5.7. MS/MS spectrum of the MH^+ ion of the 792.4 m/z peptide acquired on the "Q-ToF" instrument

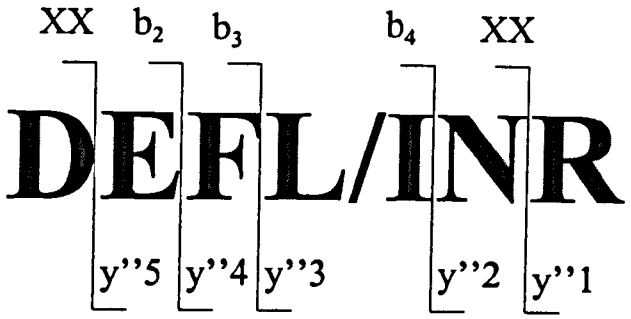


Figure 5.8. y'' and b fragment ions produced by the MS/MS experiment on the $[M+H]^+$ ion of the T21 tryptic digest peptide.

5.3 (e) Reduction and carboxyamidomethylation of the β-subunit

In order to support further the revised sequence of the T21 peptide a study of the cysteine residues in the β-subunit was undertaken. Gene sequencing predicts the presence of two cysteine residues in the β-subunit, one at position 144 and the other at position 301. The cysteine residues should be present in tryptic peptides T21 and T35. Clearly if the cysteine residue could be shown not to be present in the T21 peptide this would support the revised sequence for this peptide which does not include a cysteine residue.

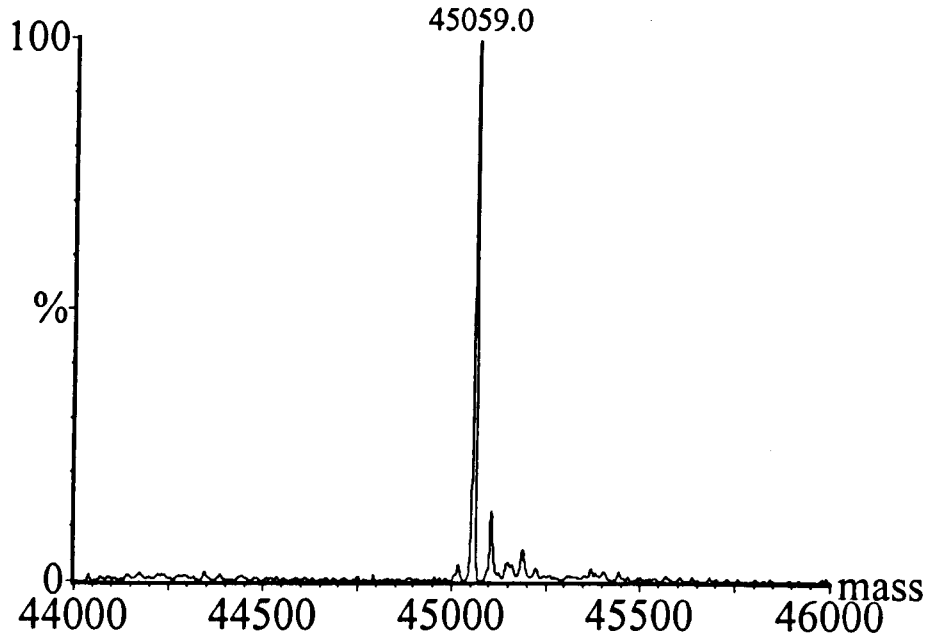


Figure 5.9. "MaxEnt" mass spectrum of the reduced and carboxyamidomethylated β-subunit

Figure 5.9. shows the "MaxEnt" deconvoluted mass spectrum of the reduced and carboxyamidomethylated β-subunit. For each of the cysteines reduced and carboxyamidomethylated (carboxyamidomethylated cysteine = $C_5H_8N_2O_2S$) an increase of 57.1 Da (average value 57.1 Da, monoisotopic value 57.02 Da) would be recorded.

The measured RMM is 45059.0 Da ±1.9 Da, which represents an increase of 57.1 Da over the measured RMM for the native β-subunit. This result suggests that there is only one cysteine residue present in the β-subunit, not two (at positions 144 and 301) as predicted by gene sequencing. In order to locate the single cysteine residue in either the

T21 or T35 digest peptide the reduced and carboxyamidomethylated β-subunit was subjected to a tryptic digestion.

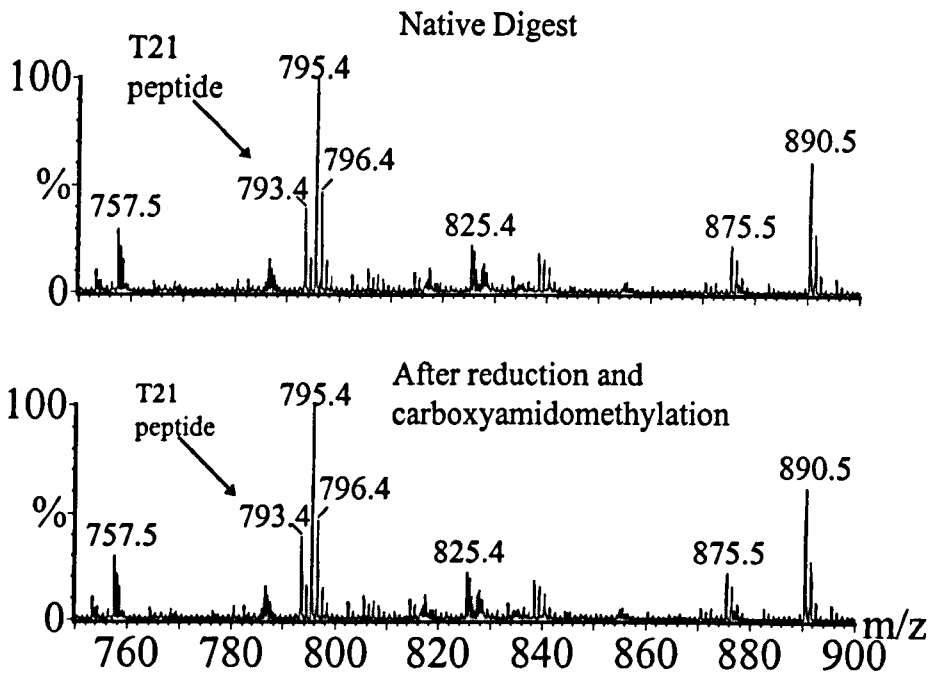


Figure 5.10. ESI spectra of a selected m/z range of the tryptic digest before and after reduction and carboxyamidomethylation.

The spectra of the tryptic digest of the native and reduced and carboxyamidomethylated β–subunit in Figure 5.10. illustrates that there is no change in the MH^+ ion of the 792.4 m/z ion upon reduction and carboxyamidomethylation. This result confirms that there is no cysteine residue present in the T21 tryptic peptide, supporting the revised sequence of DEFL/INR.

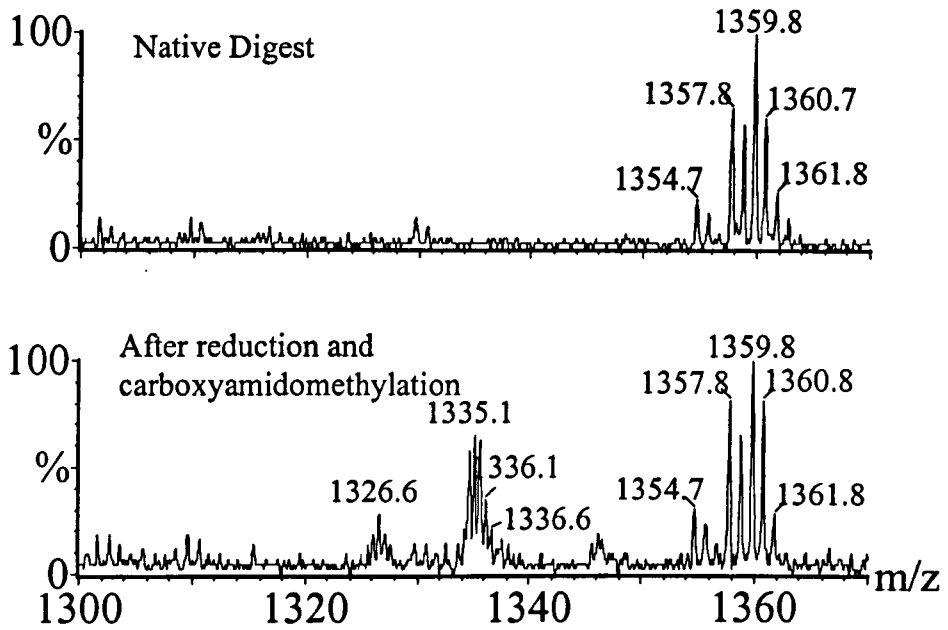


Figure 5.11. ESI spectra of a selected m/z range of the tryptic digest of the β-subunit before and after reduction and carboxyamidomethylation.

Conformation that the only cysteine residue present in the β-subunit was located in tryptic peptide T35 (position 301) was not possible from the digestion of the native β-subunit as this peptide was missing from the peptide map. Figure 5.11 shows that after reduction and carboxyamidomethylation and then a tryptic digestion a doubly charged ion of m/z 1334.60 appears which corresponds to the carboxyamidomethylation of one cysteine residue in the T35 digest peptide (T35 + carboxyamidomethylated cysteine RMM = 2667.12 Da, MH^{2+} 1334.57 m/z).

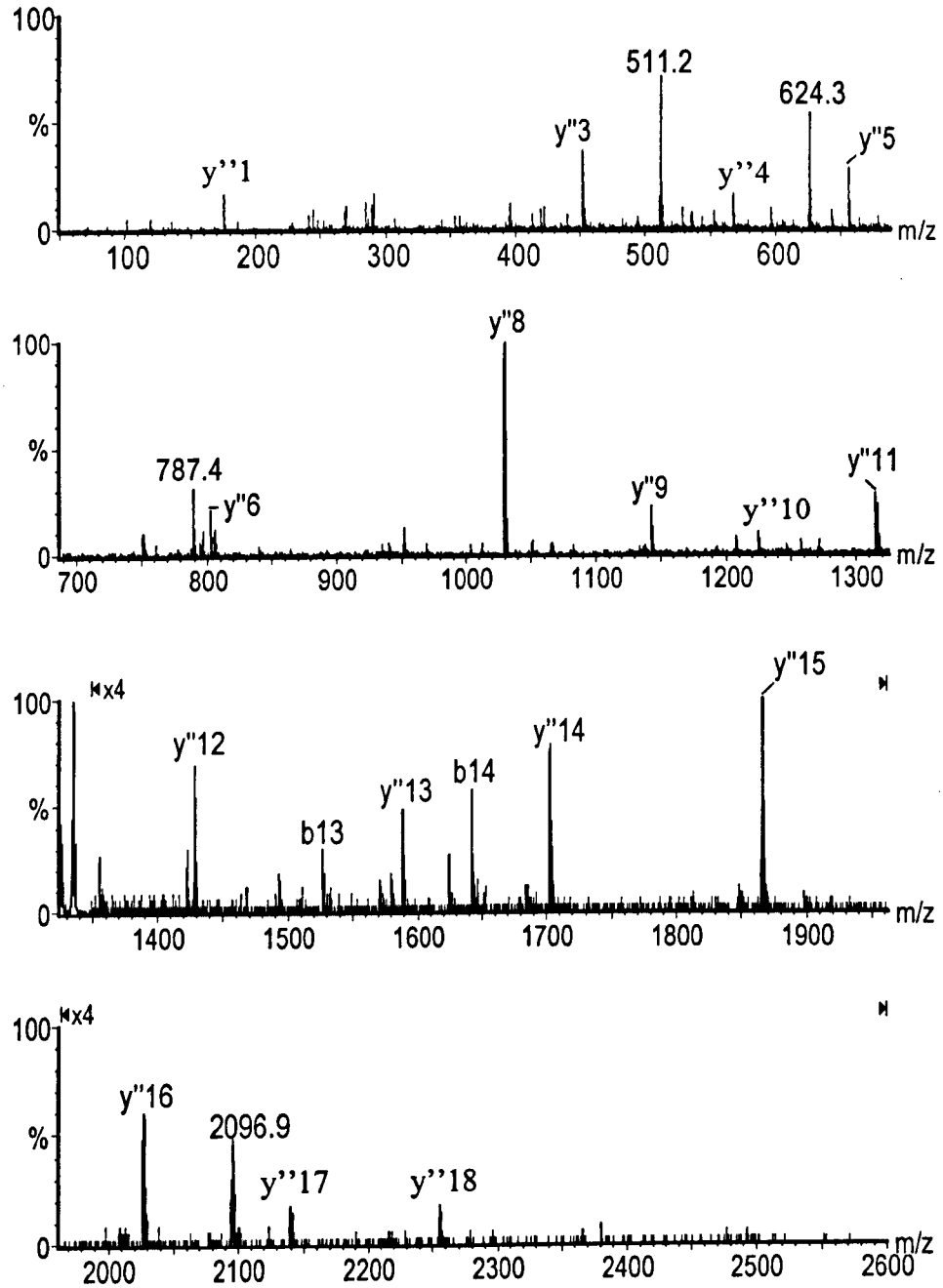
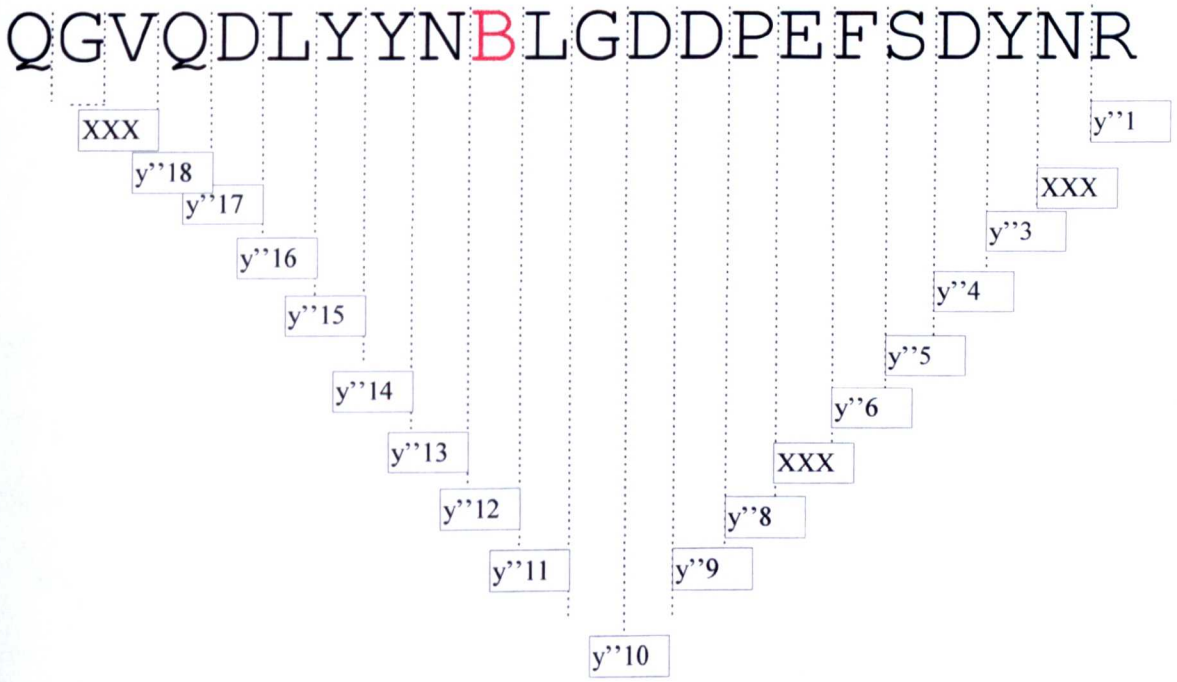


Figure 5.12. MS/MS spectrum of the 1334.6-1336.1 m/z cluster of isotope peaks from the reduced and carboxyamidomethylated tryptic digestion of the β-subunit.

MS/MS analysis, shown in Figure 5.12., of the 1334.6-1336.1 m/z cluster of isotope peaks, undertaken on the "Q-ToF" instrument, produced a partial product ion spectrum that was consistent with that expected of the T35 peptide containing a carboxyamidomethylated cysteine. Thus the one cysteine present in the β-subunit is located in the T35 digest peptide and not in the T21 peptide which carries the sequence error.



B = Carboxyamidomethylated cysteine

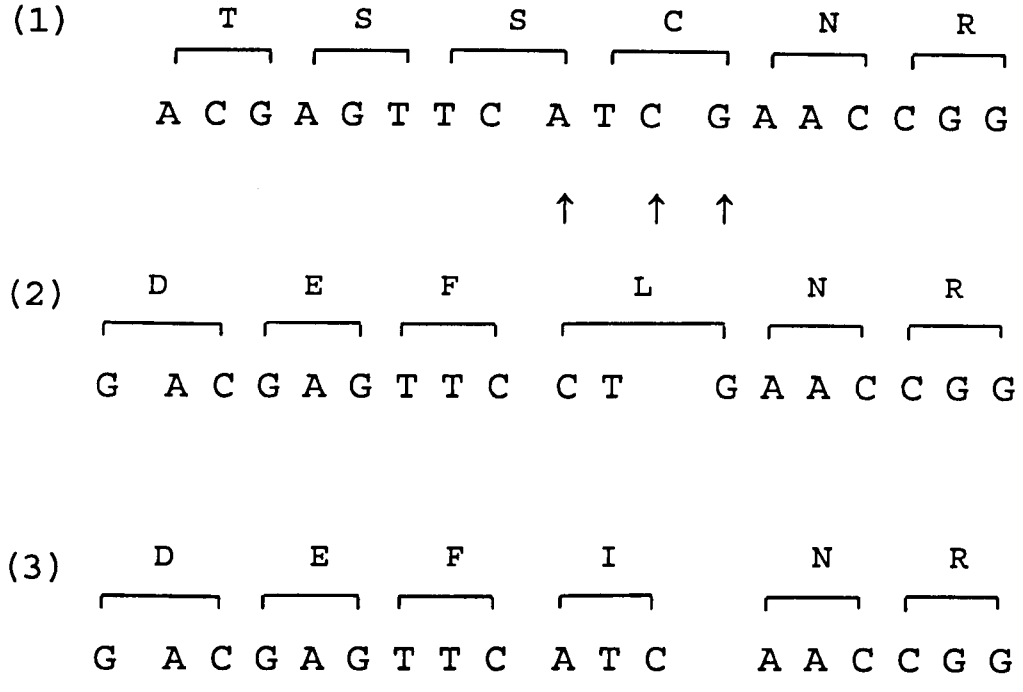
Figure 5.13. y'' fragment ions produced by the MS/MS experiment of the T35 tryptic digest peptide after reduction and carboxyamidomethylation.

5.4 Conclusions

ESI-MS analysis showed that the measured RMM of the β-subunit disagreed with the value calculated on the basis on the gene sequence. A tryptic digest of the native β-subunit was performed in order to locate the discrepancy of 126.1 Da. The digest confirmed the majority of the primary sequence (90%). The digest peptide T21 (residues 141-146) was identified as the location of the +126.1 Da. MS/MS analysis of the MH^+ ion of the T21 peptide suggested a revised sequence of DEFL/INR for the peptide which fully accounts for the difference in mass. Protein and peptide database searches ruled out the possibility that the 792.4 Da peptide was a product of auto-proteolysis of the bovine pancreatic trypsin reagent used.

Reduction and carboxyamidomethylation of the subunit confirmed that there was only one cysteine in the protein (as opposed to the two predicted by gene sequencing), and that this was present in the T35 peptide and not in the T21 peptide. The tryptic digest of the reduced and alkylated β-subunit confirmed 95% of the primary sequence.

Further support for the revised sequence comes from the fact that the DNA nucleotide sequence, which originally suggested the amino acid sequence to be TSSCNR, (1), has to be changed only very little to give rise to the two possible revised sequences, DEFLNR (2) and DEFINR (3):



in which the codes have been displaced slightly to emphasise the similarity. The two possible revised sequences therefore can arise from a single base change indicated by the arrows and account satisfactorily for the error in mass of the b-subunit and each is consistent with all the available evidence.

This chapter describes how the use of ESI/MS correctly identified and corrected errors arising from a slight mis-reading of gene-sequence data to which attention was first drawn by an X-ray analysis of the protein. Sensitivity problems associated with the use of conventional CNBr cleavage followed by separation by HPLC and tryptic digestion

of any peptide of interest were largely overcome by the use of a tryptic digest of the whole protein followed by the use of CID to analyse the appropriate tryptic peptide without chromatographic separation although the use of the more sensitive "Q-ToF" instrument was required to obtain the complete sequence unequivocally.

5.5 References

1. J. Colby, D. I. Stirling, H. Dalton, *Biochem. J.*, **165**, 395 (1977).

2. D. H. Ehhalt, 'In *Microbial Production and Utilization of Gases'* H. G. Schlegel, G. Gottschalk and N. Pfennig,(Eds.), Goltze, Göttingen. (1976), pp 13.

3. H. Dalton, B. Golding, B. Waters, R. Higgins, J. A. Taylor, *J. Chem. Soc., Chem. Comm.*, 482 (1981).

4. D.J. Leak, H. Dalton, *Biocatalysis*, **1**, 23 (1987).

5. J. Green, H. Dalton, *J. Biol. Chem.*, **264**, 17698 (1989).

6. R. Oldenhuis, D. B. Janssen, In '*Microbial Growth on C1 Compounds'*, J. C. Murrell and D. P. Kelly (Eds.) Intercept, Andover, 1993 pp.121.

7. J. A. Zahn, A. D. Dispirito, *J. Bact.*, **178**, 1018 (1996).

8. A. K. Sheimke, S. A. Cook, T. Miley, P. Singleton, *Arch. Biochem. Biophys.*, **321**, 421 (1995)

9. K. E. Liu, S. J. Lippard, *Adv. Inorg. Chem.*, **42**, 263 (1995)

10. J. D. Lipscomb, *Ann. Rev. Microbiol,* **48**, 371 (1994)

11. A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, P. Nordlund, *Nature*, **366**, 537 (1993).

12. A. C. Stainthorpe, J. C. Murrell, G. P. C. Salmond, H. Dalton, V. Lees, *Arch. Microbiol.*, **152**, 154 (1989).

13. A, Buzy, A. L. Millar, V. Legros, P. C. Wilkins, H. Dalton, K. R . Jennings,. *Eur. J. Biochem.* **254**, 602 (1998).

14. S. J. Pilkington and H. Dalton, 'In Methods in Enzymology', M. E. Liddtrom, (Ed) Academic press: San Diego (1990), *188*, pp. 181.

Chapter 6.

Epitope-Paratope Interactions Between HIV-1 And A 17 Amino Acid Micro-Antibody

6.1 Introduction

Animals can produce specialised proteins called antibodies in response to the presence of particular antigens (a substance that the immune system regards as foreign or potentially dangerous) and render it harmless. The precise region of an antigen to which an antibody binds to is called an epitope. Epitopes can be classified as being either linear or non-linear. Linear epitopes are composed of a continuous chain of amino acid residues, whilst non-linear epitopes are made up of two or more short chains of amino acid residues which are not directly linked but are topographically localised by the antigen's tertiary structure.

The basic structural unit of an antibody molecule consists of four polypeptide chains, two identical light (L) chains and two identical heavy (H) chains. The four chains are held together by a combination of non-covalent and covalent (disufide) bonds. The molecule is composed of two halves, each with the same epitope binding site, and both L and H chains usually co-operate to form the antigen-binding surface.

Each heavy and light chain has a variable amino acid residue region (V_L & V_H). Analysis of the amino acid sequence of V_L and V_H has shown regions of hypervaribility. Three such regions are present in each heavy and light chains. It is these areas of hypervaribility that form the precise epitope-binding site of the antibody molecule. As the epitope-binding site is complementary to the structure of the epitope, the hypervariable regions are also called complementarily determining regions (CDRs).

The CDRs form what is called the paratope of the antibody, which binds to the epitope of the antigen.

The mechanisms of action of neutralising antibodies that bind to viruses which enter the body are complex and diverse [1,2].

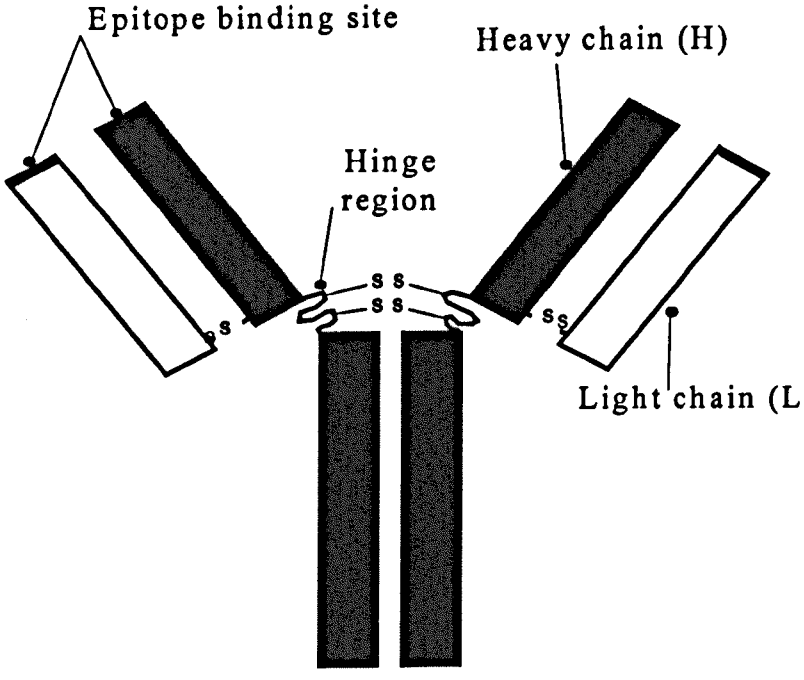


Figure 6.1. Basic structure of an antibody molecule

In vitro there are three classes of anti-human immunodeficiency virus type 1 (HIV-1) antibody action: (I) inhibition of virus attachment to target cell CD4 receptors [3-6]; (II) inhibition of the viral genome and associated proteins into the cell interior [4,7] and (III) inhibition of a subsequent event beyond entry [4,7,8] which has not yet been defined.

Human immunodeficiency virus type 1 (HIV-1) [9] is the causative agent of acquired immunodeficiency syndrome (AIDS). To date, HIV-1 infects 30 million people world wide, an estimate that has exceeded predictions of its rate of spread (UN World AIDS Day report; 26th Nov. 1997). The most cost effective strategy to deal with this global epidemic is the production of a vaccine that will stimulate an effective cellular and antibody immune response. Present approaches, utilising antibodies, have failed to consistently stimulate antibodies that are effective against diverse primary strains of HIV-1 [10] although examples of antibodies with high neutralisation efficiency have been isolated (b12 [11], 2G12 [12] and 2F5[13]).

Many different methods have been used in the investigation of paratope-epitope interactions. X-ray crystallography and NMR of FAB-virus or FAB-peptide antigen crystals (a FAB fragment is a product of an enzymatic digest of an antibody) have provided detailed structural information in several cases, although both are very time consuming and require a large amount of sample.

Alternative methods include solid-phase immunoassay where an anti-body specific for the protein of interest is attached to a polymermic support. Upon addition of the sample

antigen a complex between the two is formed. A second antibody specific for a different site of the antigen is then added. The second antibody carries a radioactive or fluorescent label so that it can be detected with high sensitivity. The amount of second anti-body is proportional to the antigen in the sample. The sensitivity of the assay can be greatly increased if the second antibody is attached to an enzyme, such as alkaline phosphatase. This enzyme can rapidly convert many molecules of an added colourless substrate into coloured products. Such an assay is termed an enzyme-linked immunsorbent assay (ELISA).

An alternative method is the use of mass spectrometry. In 1990, Suckau *et al* [14] combined immunoassay and ^{252}CF plasma desorption mass spectrometry to identify epitopes to the antibody h453. They compared the enzymatic digest patterns of the free antigen with those obtained when the anitgen was complexed with the antibody. This method took advantage of the fact that, when complexed, the antigen was protected from digestion by the antibody and would thus produce a different digest pattern from those obtained from the free antigen.

This approach to epitope identification, sometimes referred to as epitope excision, has the advantage of being suitable for the determination of both linear and non-linear epitopes. Papac *et al* [15] used a similar method of determining the epitope to an immobilised antobombesin antibody but utilised matrix assisted laser desorption ionisation (MALDI)-MS. Parker *et al* [16] have also employed the epitope excision

approach to determine an epitope on the HIV $_{IIIB}$ Gag protein again using MALDI-MS as have the majority of subsequent studies of this type.

An alternative method of epitope identification is the epitope extraction approach. This involves the pre-digestion of the antigen before reaction with the antibody. Those digest peptides that contain the epitope are selectively bound by the antibody, leaving the non-epitope peptides to be removed. Zhao *et al* [17] have identified a linear epitope in the peptide melitten using this strategy. Jeyarajah *et al* [18] have demonstrated that the bound epitope containing peptides need not be eluted from the immobilised antibody prior to MALDI -MS analysis. Whilst Madit *et al* [19] have shown that either epitope excision or extraction may be performed without immobilising the antibody. In these experiments separation of the bound and unbound peptides was performed by means of ultrafiltration. Lyubarskaya *et al* [20] have employed affinity capillary electrophoresis with electrospray ionisation mass spectrometry for the screening of digest peptides containing epitopes.

In a recent study Lu *et al* [21] have demonstrated the identification of protein antigens in complex mixtures with a magnetic bead monoclonal antibody.

In the majority of the MALDI-MS studies performed so far the antibody has had to be immobilised before being incubated with the complete or digested antigens. The non-epitope peptides then have to be washed away, before the epitope-angtigen complex can submitted for mass spectrometric analysis. Both of these steps increase the complexity of the process and the time required for the analysis.

Another disadvantage of these types of analyses is that as the antibody is covalently bound to the solid matrix , only the non-covantly bound peptides are desorbed during the MALDI process. The actual complex between the antibody and the epitope is not observed by mass spectrometry. This can increase the possibility of mis-interpretation of non-specific interactions between the antibody and those digest peptides of the antigen that do not contain the epitope.

The work detailed in this chapter is concerned with the binding activities of a micro-anitbody (Micro-Ab). The Micro-Ab is a 17 amino acid cyclised peptide derived from the third heavy chain CDR of a murine monoclonal antibody (MAb) IgG1, called F58 [22]. MAb F58 neutralises several T-cell lab-adapted (TCLA) and primary strains of HIV-1. The Micro-Ab was shown previously to neutralise five TCLA strains of HIV-1 [8,13] and act by inhibiting an event beyond the fusion of HIV-1 with its target cell membrane.

MAb F58 binds to a region on the glycoprotein (gp) spikes present on the surface of the HIV-1 lipid envelope membrane [23-24]. There are about 70 spikes on the surface of each viron [25]. The spikes consist of a transmembrane glycoprotein gp41 molecule, interacting non-covalently with a glycoprotein gp120 molecule, to form an oligometric structure which recent data proposes is a trimer [26,27,28]. Gp120 houses the CD4 primary cell receptor binding site (or CD4-binding site) [29]. It is believed that Gp120 and gp41 are involved in the complex interactions between HIV-1 and its target cell, that result in HIV-1 genome entry into the cell interior.

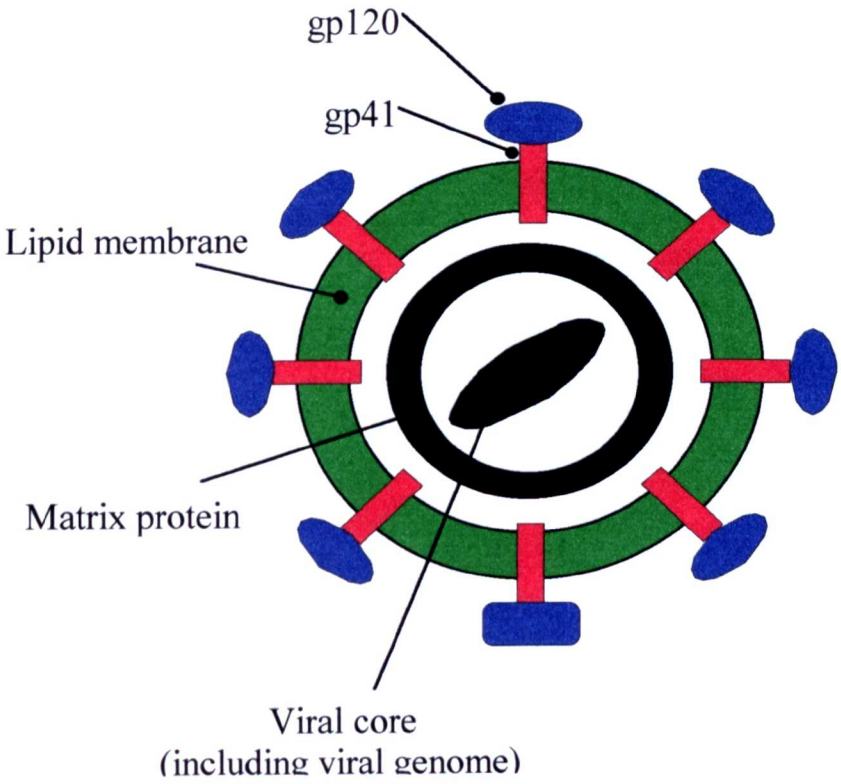


Figure 6.2. Schematic representation of HIV-1

Gp120 is highly glycosylated and composed of regions demonstrating degrees of sequence variability between different strains of HIV-1. In one monomeric gp120 molecule there are five conserved regions (C1-C5) and five variable regions (V1-V5). Neutralisation epitopes in gp120 have been mainly identified in the CD4 binding site,

V3 and V2 regions. It is to the third variable region (V3) that MAb F58 specifically binds.

At the tip of the V3 region, the minimum epitope for F58 binding is IxxGPGR as determined by ELISA studies. The strongest binding, however, was seen with a peptide representing the amino acid sequence RKSIRQRGPGRA. Previous studies have shown that Micro-Ab competes with MAb F58 for binding to V3 peptides representing this region and suggest the Micro-Ab binds to this region of gp120.

A: CDLIYYDYEEDYYFDYC

B: RKSISIGPGRAFFATGD

C:CTRPNNNTRKSIHIGPGRAFYATGEIIGDIRQAHC

Figure 6.3. Amino acid sequences of; A: Micro-Ab; B:V3 Z-321 peptide; C:V3 BRA peptide

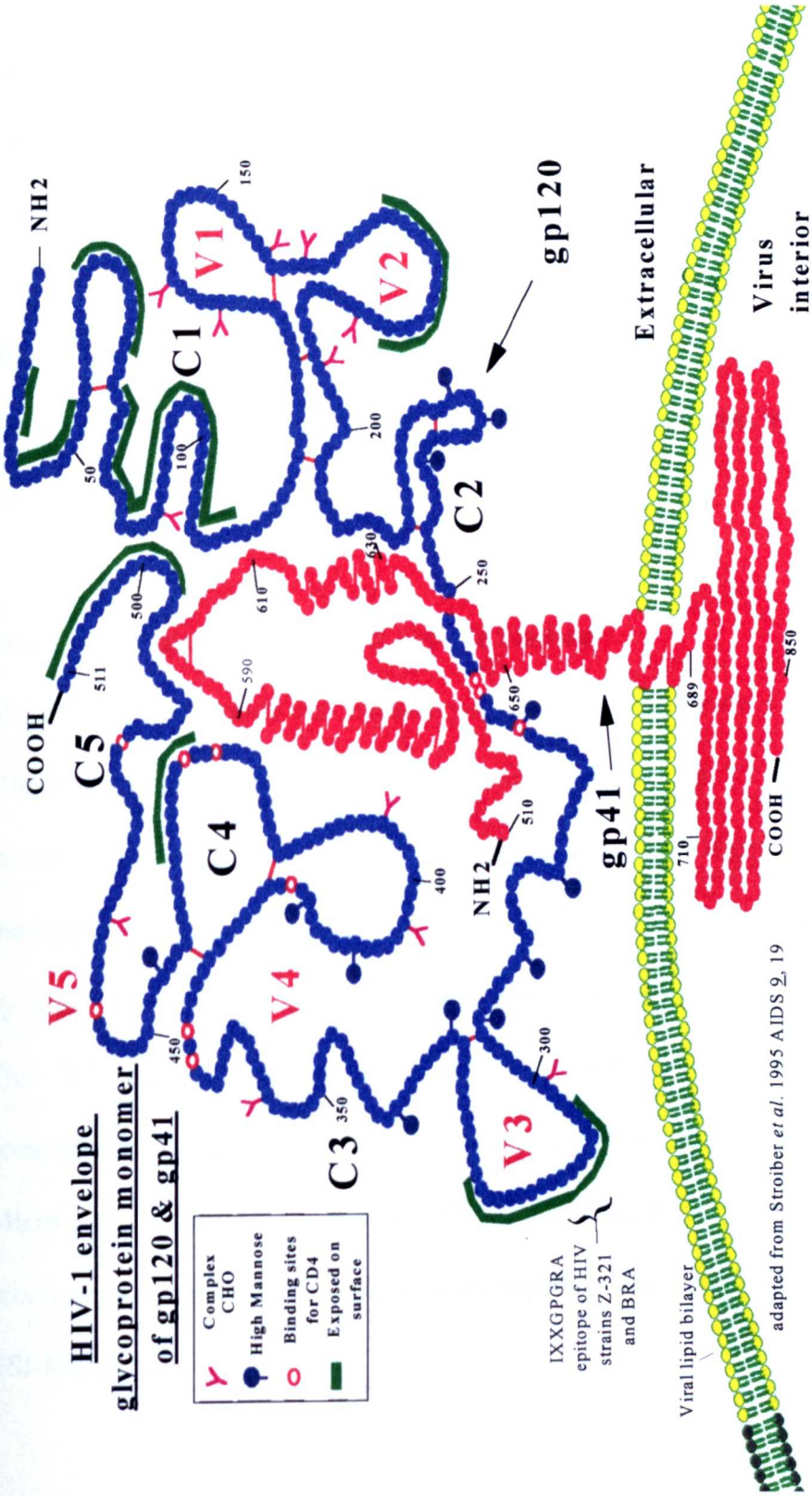


Figure 6.4. Schematic of the gp 41 and gp 120 of the HIV-1 envelope protein.

6.2. Experimental

Sample preparation

Lyophilised micro-antibody, HIV-1 Z-321 V3 peptide (MRC-ADP; ARP 765) and HIV-1 BRA V3 peptide (MRC-ADP; EVA 7026) were separately dissolved in 50 mM ammonium acetate at 1mg/mL.

Micro-Ab and V3 peptide binding

The Micro-Ab and V3 peptides, present in equimolar quantities, were incubated at 37°C for 1 hour.

Proteolytic digestion and digest peptide binding

Z-321 or BRA peptide and chymotrypsin/tryspin in the ratio 1:50 were dissolved in 50mM ammonium acetate. For partial digestion of the V3 peptides the digestion mixture was incubated at 4°C for 2 minutes; for complete digestion of the V3 peptides the digestion mixture was incubated for 1 hour at 37°C. The digest reagent was removed by ultrafiltration using a Microcon microconcentrator (Amicon, INC. Beverley MA). The V3 digest peptides were either diluted to working concentrations in acetonitrile/water plus 0.1% formic acid for ESI/MS analysis, or incubated with the Micro-Ab (1:1 ratio) for 1 hour at 37°C to allow any binding to occur. Digest peptide mixtures or digest peptide incubated with the Micro-Ab were then submitted directly for ESI-MS analysis.

Mass Spectrometry

All experiments were performed in a "Quattro II" tandem mass spectrometer of QhQ configuration equipped with electrospray ionisation. Instrumental conditions were as described in chapter 3. Micro-Ab and V3 peptides were dissolved in CH_3CN/H_2O plus 0.1% HCOOH at a concentration of approximately 3pm/μL Calibration was performed using the multiply charged ions of horse heart myoglobin or the monoisotopic ions of NaI.0

6.3. Results and discussion

6.3 (a) Relative molecular mass determination

Figure 6.5. shows the ESI spectrum of the Micro-Ab sprayed in a solvent consisting of a 1:1mixture of H_2O/CH_3CN plus 0.1% HCOOH. Peaks due to a singly and a doubly charged ion corresponding to a peptide of measured monoisotopic RMM 2291.7 Da ± 0.2 Da are present. This compares well with the calculated value of 2291.8 Da for the Micro-Ab, confirming both its primary sequence and the presence of one disulfide bond in the peptide. Na^+ adducts are also present in the spectrum.

The ESI spectrum of the HIV-1 V3 Z-321 peptide is shown in Figure 6.6. The spectrum shows peaks assigned as singly, doubly and triply charged ions corresponding to a peptide of monoisotopic RMM 1778.9 Da ± 0.08 Da, this is in exact agreement with the calculated RMM and therefore confirms the primary sequence of the V3 Z-321 peptide.

The ESI spectrum of the larger HIV-1 V3 BRA peptide is shown in Figure 6.7. The spectrum shows a series of peaks corresponding to ions of the charge states 2+ to 6+.

The measured average RMM of the V3 BRA peptide is 3866.5 Da ± 0.4 Da. Again this compares very well with the calculated RMM of the peptide of 3866.4 Da and confirms the primary sequence of the HIV-1 V3 BRA peptide and the presence of one disulfide bond. An unidentified peptide of RMM 3780 Da is also present.

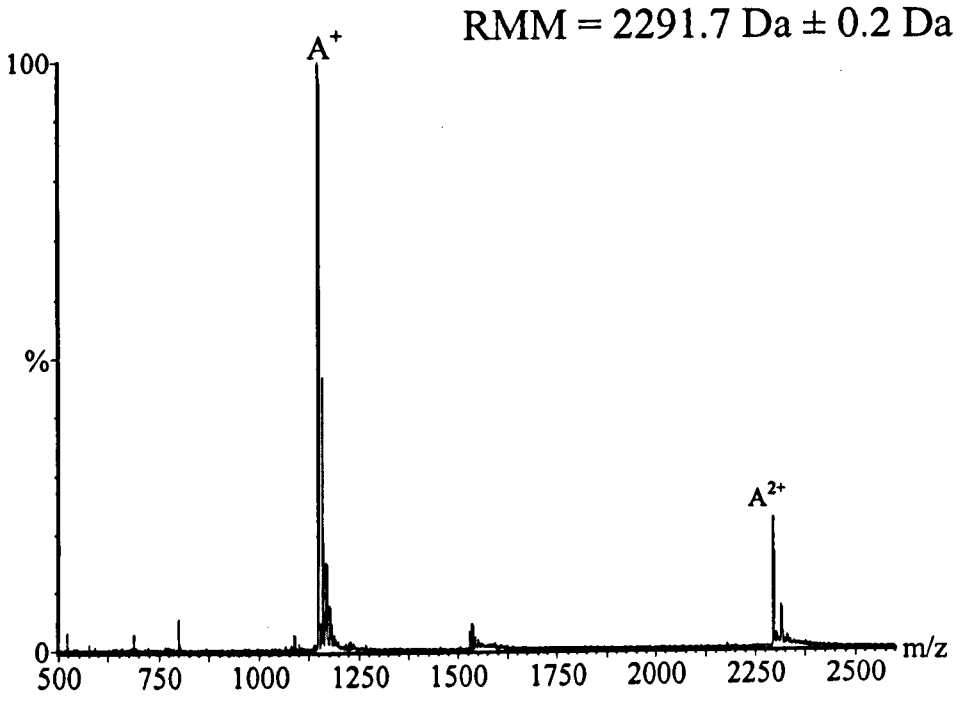


Figure 6.5. ESI spectrum of the Micro-Ab

The measured RMMs for the Micro-Ab and the V3 Z-321 and BRA peptides were in good agreement with their expected values and therefore confirm their primary sequences. The presence of a disulfide bond in both the Micro-Ab and the BRA peptide is confirmed as, in each case, the RMM would be 2 Da lower without the bond.

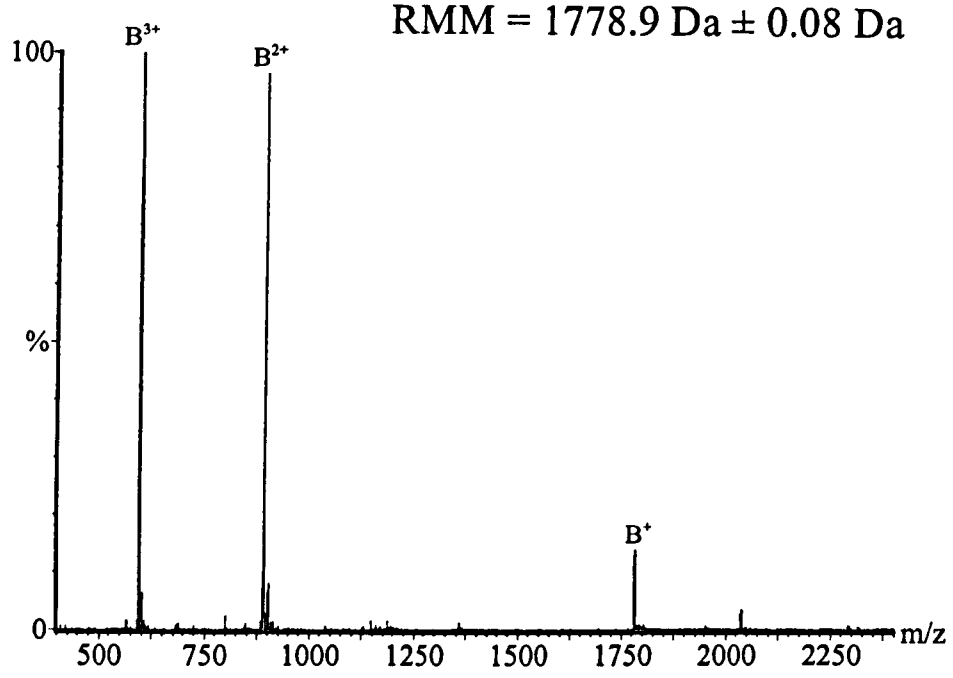


Figure 6.6. ESI spectrum of the HIV-1 V3 Z-321 peptide

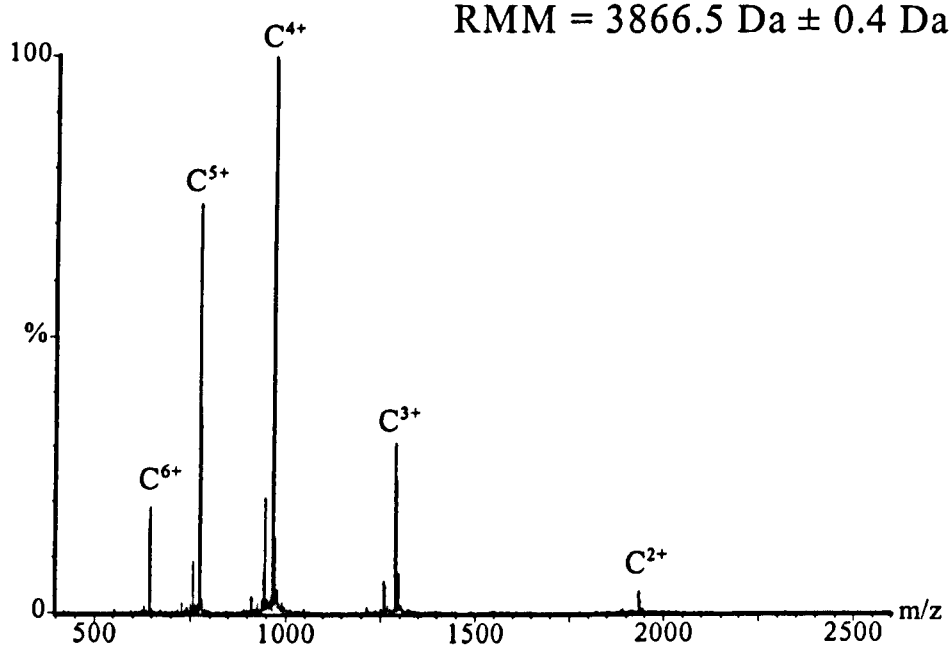


Figure 6.7. ESI spectrum of the HIV-1 V3 BRA peptide.

6.3 (b) Incubation of the Micro-Ab with the V3 peptides

To determine whether or not they interacted, the Micro-Ab and the V3 peptides were incubated using the conditions that resulted in 95% loss of infectivity of HIV-1 strain.

The ESI spectrum of the Micro-Ab incubated with the V3 Z-321 peptide is shown in Figure 6.8.

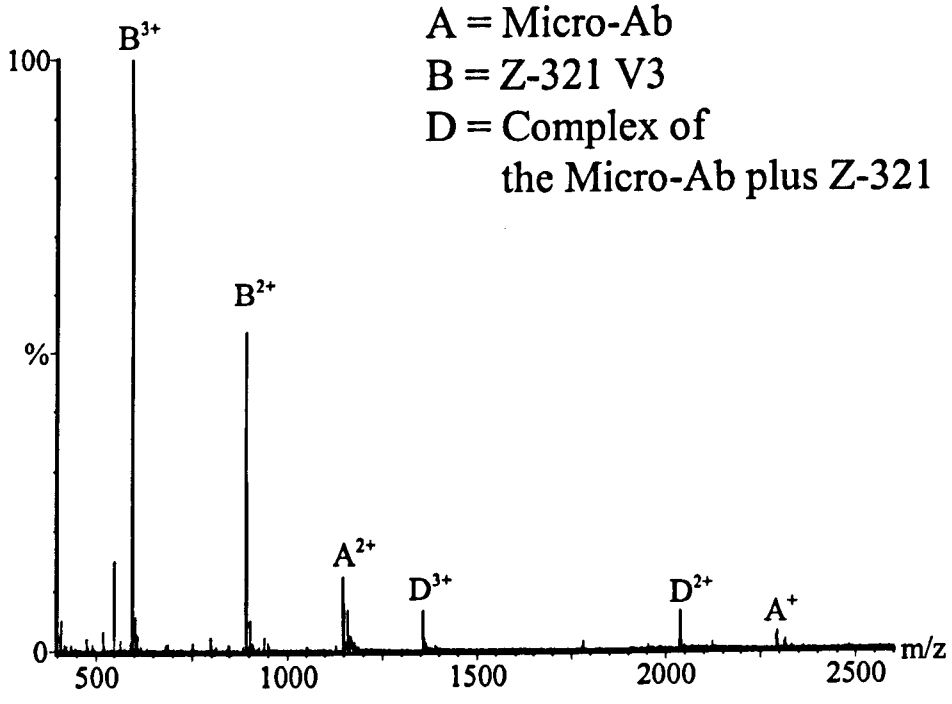


Figure 6.8. ESI spectrum of the Micro-Ab incubated with the V3 Z-321 peptide at eqimolar concentrations for 1 hour.

The spectrum shows the typical contributions of the Micro-Ab (labelled A) and the V3 Z-321 peptide (labelled B). A third series of peaks is also present (labelled D) these are assigned as a doubly and a triply charged ion. These ions correspond to a protein of 4072.7 Da; this value compares well with the expected average RMM of a complex between the Micro-Ab and the V3 Z-321 peptide of 4073.4 Da. This result indicates that the complex formed between the Micro-Ab and the V3 Z-321 peptide in solution by a non-covalent interaction can be maintained and detected in the gas phase.

The above procedure was repeated with the Micro-Ab and the V3 BRA peptide. Figure 6.9. shows the spectrum of the products of the incubation of the Micro-Ab and the V3 BRA peptide.

Again the spectrum shows the typical contributions of the Micro-Ab (labelled A) and the V3 BRA peptide (labelled C). An additional series of peaks are also present, labelled E, these are assigned as triply and quadruply charged ions, which correspond to a protein of average RMM 6159.1 Da. This compares favourably with the expected RMM of a complex between the Micro-Ab and the V3 BRA peptide. These results, as for the Micro-Ab and the V3 Z-321, suggest that the interaction between the Micro-Ab and the V3 BRA peptide can be preserved by ESI and thus detected directly by mass spectrometry.

The results of the relative molecular mass determination and incubation studies suggest that the low RMM of the Micro-Ab makes it more amenable to ESI-MS studies than an

investigation using a standard antibody (typical RMM approximately 150 kDa). The observation that the non-covalent interaction between the Micro-Ab and the V3 peptides can be maintained by electrospray ionisation and thus detected directly by mass spectrometry means the immobilisation of the antibody and the subsequent washing procedures characterised by the majority of the previous mass spectrometric analysis found in the literature can be omitted.

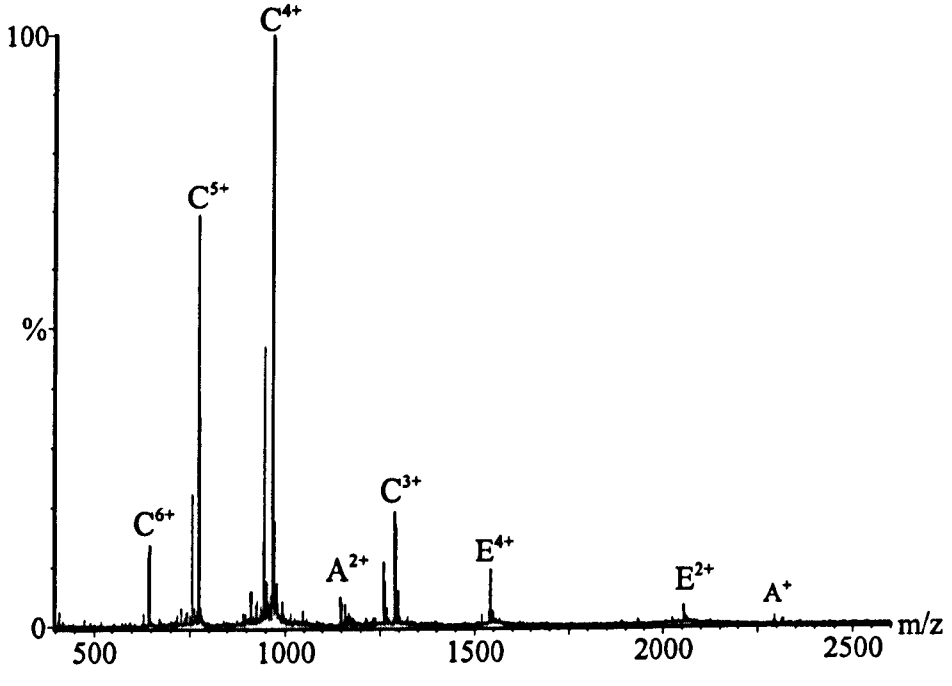


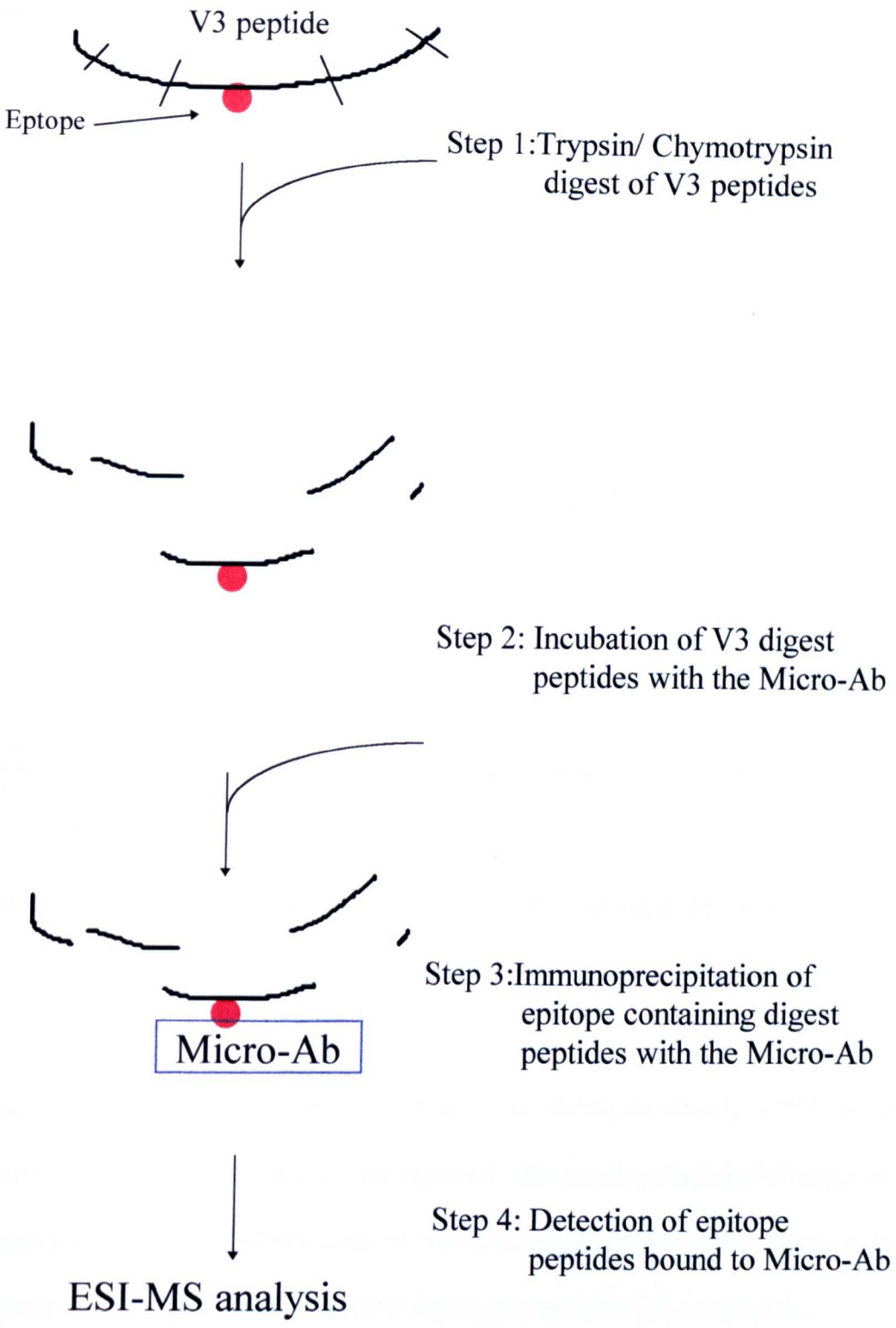
Figure 6.9. ESI spectrum of the Micro-Ab incubated with the V3 BRA peptide

6.3 (c) Epitope identification

As outlined in the introduction at the beginning of the chapter two main methods for epitope identification exist; epitope excision where the antibody-antigen complex is formed and then subjected to digestion, or the epitope extraction method where the antigen is predigested and the digest peptides (only some of which contain the epitope) are incubated with the antibody.

In order to use the epitope excision method it was considered necessary to separate the unbound Micro-Ab and V3 peptide from the Micro-Ab-V3 complex in solution so that the digestion pattern of the bound and free antigen could be compared.

Previous efforts to use the epitope excision method without separating the free Micro-Ab and respective V3 peptide from the bound complex indicted that the digestion patterns were difficult to assign unambiguously. All attempts to perform this separation by HPLC and FPLC were unsuccessful. The epitope extraction method was therefore employed, where the antigen would be predigested before being incubated with the Micro-Ab (Scheme 6.1.).



Scheme 6.1. Schematic representation of the epitope identification experiment

The ESI spectrum of the chymotryptic digestion of the Z-321 V3 peptide is shown in Figure 6.10.

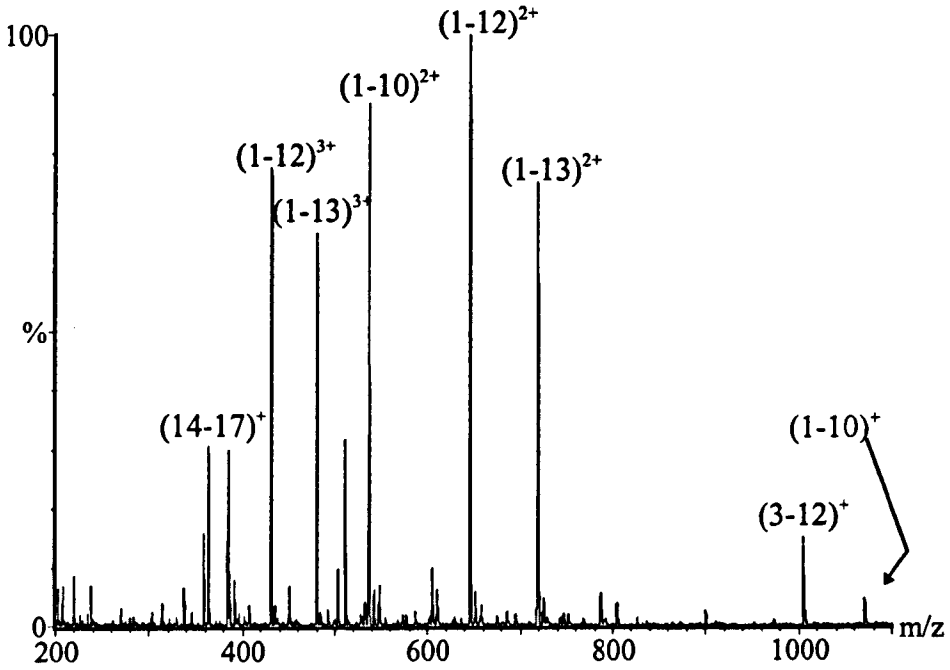


Figure 6.10. ESI spectrum of the products of a chymotryptic digest of the V3 Z-321 peptide

Using the "Biolynx" protein/peptide database the predicted products of a chymotryptic digestion of the V3 Z-321 peptide were obtained. The good agreement between the measured and predicted RMM values allowed the majority of the peaks present in the spectrum to be assigned as chymotryptic digest ions of the V3 Z-321 peptide.

The major digestion products were found to be the digest peptides consisting of the 1-10, 1-12 and 1-13 amino acid residues. All the chymotryptic digest peptides produced

are listed in table 6.1. The next step in the procedure was to incubate the digest peptides with the Micro-Ab for immunoprecipitation of the epitope containing digest peptides to occur.

Z-321 peptide Fragment	Charge State	Expected RMM (Da)	Measured RMM (Da)	Difference (Da)
1-10	2+	535.8	535.8	0.0
1-12	2+	644.9	644.9	0.0
1-12	3+	430.2	430.2	0.0
1-13	2+	718.4	718.5	0.1
1-13	3+	479.3	479.3	0.0
14-17	1+	363.2	363.3	0.1

Table 6.1. Expected and measured RMM values of the chymotryptic digest products of the V3 Z-321 peptide

The ESI spectrum of the chymotryptic digest products of the Z-321 peptide incubated with the Micro-Ab is shown in Figure 6.11. As in Figure 6.10. the most abundant peaks are assigned as ions corresponding to the expected digest peptides of the V3 Z-321 peptide. Also present are peaks assigned as the characteristic singly and doubly charged ions of the Micro-Ab.

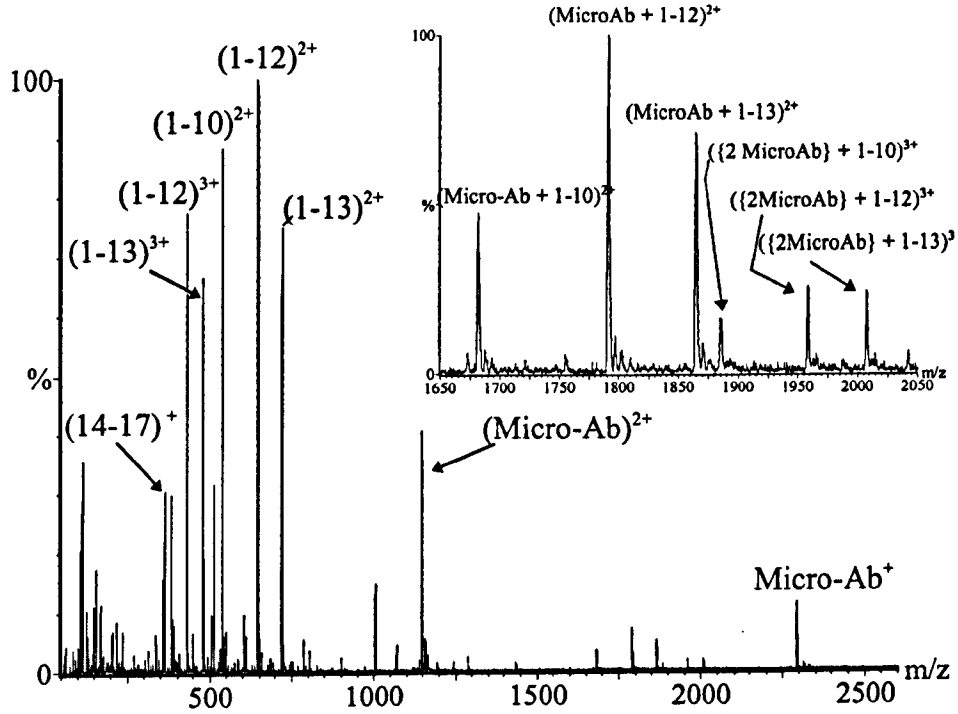


Figure 6.11. ESI spectrum of the chymotryptic digest products of the Z-321 peptide incubated with the Micro-Ab. Inset shows expanded m/z range.

The inset of Figure 6.11. shows an expanded view of the m/z range 1650-2050. The peaks contained within this range were assigned as ions corresponding to the Micro-Ab binding to the digest peptide consisting of amino acid residues 1-10, the digest peptide consisting of amino acid residues 1-12 and the digest peptide consisting of amino acid residues 1-13. Peaks corresponding to two Micro-Abs binding to the digest peptides consisting of amino acid residues 1-10, 1-12 and 1-13 were also present.

No binding of the Micro-Ab to the digest peptide consisting of amino acid residues 14-17 was observed , despite its presence in the digest peptide mixture. These data suggest that amino acid residues 14-17 are not required for binding of the V3 Z-321 peptide to the Micro-Ab, and that the epitope is to be found in amino acid residues 1-10.

No non-specific interactions between the Micro-Ab and other digest peptides present in the digestion mixture were observed. The digestion and binding data are summarised in Figure 6.12.

Additional information on the epitope, now believed to be contained in amino acid residues 1-10, was sought by subjecting the V3 Z-321 peptide to a tryptic digestion. Trypsin digests on the C-terminal side of arginine and lysine residues and so was expected to produce either/or both of the peptides containing amino acid residues 2-10 and 3-10. A partial and complete tryptic digestion of the Z-31 was performed to produce both the peptides consisting of the 1-10 and the 2-10 / 3-10 amino acid residues to confirm the results obtained with the chymotryptic digestion of the V3 Z-321 peptide in one experiment.

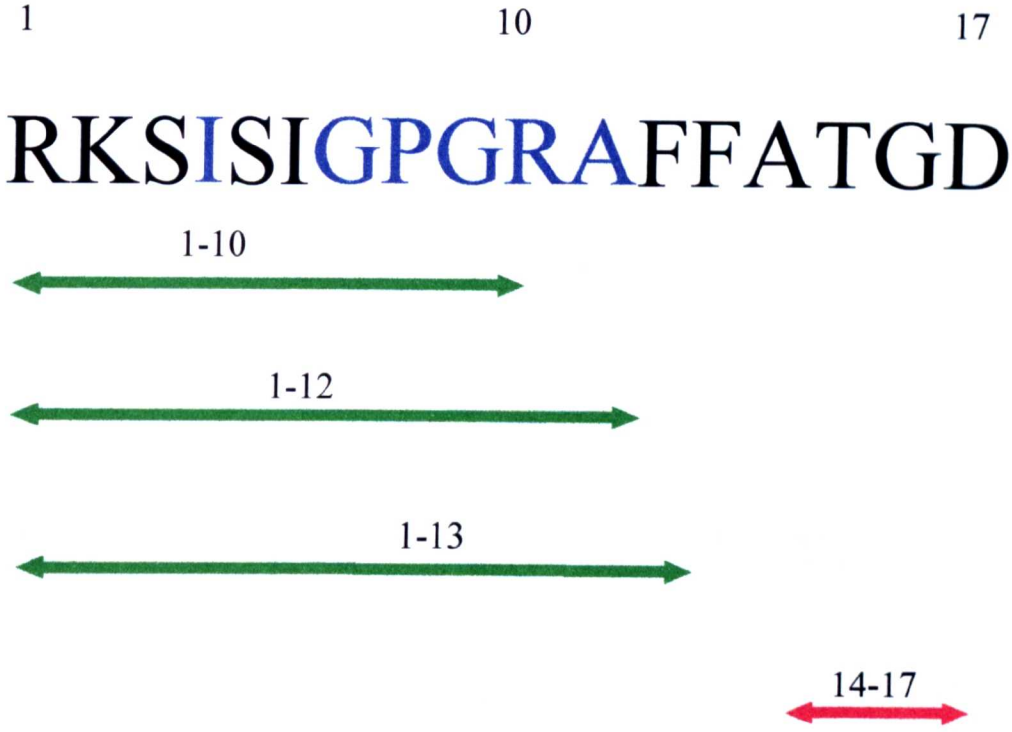


Figure 6.12. Major peptides produced by a chymotryptic digestion of the V3 Z-321 peptide. The green lines represent peptides that were bound by the Micro-Ab, the red line represents the peptide that was not bound by the Micro-Ab.

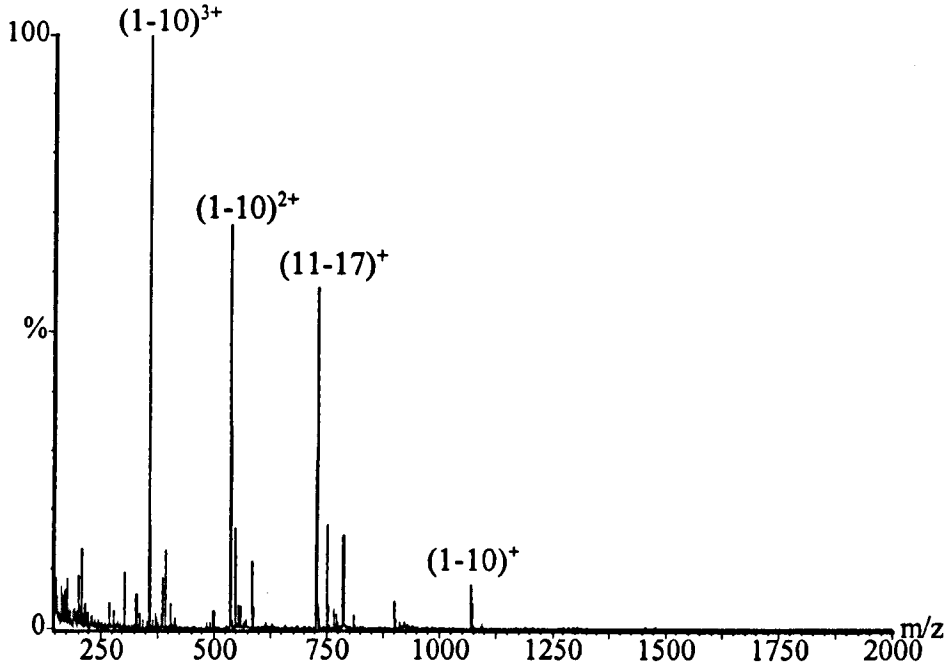


Figure 6.13. ESI spectrum of the products of a partial tryptic digestion of the V3 Z-321 peptide

The partial tryptic digest of the V3 Z-321 peptide produced the digest peptides consisting of amino acid residues 1-10 and 11-17 and a small amount of that consisting of the 3-10 amino acid residues as assigned in Figure 6.13. and illustrated in Table 6.2.

Z-321 peptide fragment	Charge state	Expected RMM (Da)	Measured RMM (Da)	Difference (Da)
1-2	1+	303.2	303.1	0.1
1-10	1+	1070.6	1070.6	0.0
1-10	2+	535.6	536.0	0.2
1-10	3+	357.6	357.6	0.0
3-10	1+	786.5	786.5	0.0
3-10	2+	393.7	393.8	0.1
11-17	1+	728.3	728.4	0.1

Table 6.2. Expected and measured RMM values for the products of a partial tryptic digest of the V3 Z-321 peptide

Figure 6.14 shows the ESI spectrum of the products of the partial tryptic digestion of the Z-321 peptide incubated with the Micro-Ab. The spectrum contains peaks which correspond to the expected digest peptides and the Micro-Ab. The inset of the spectrum shows two peaks which correspond to doubly and triply charged ions assigned as being due to the Micro-Ab binding to the digest peptide consisting of amino acid residues 1-10. No binding of the Micro-Ab to the peptide containing amino acid residues 11-17 was detected.

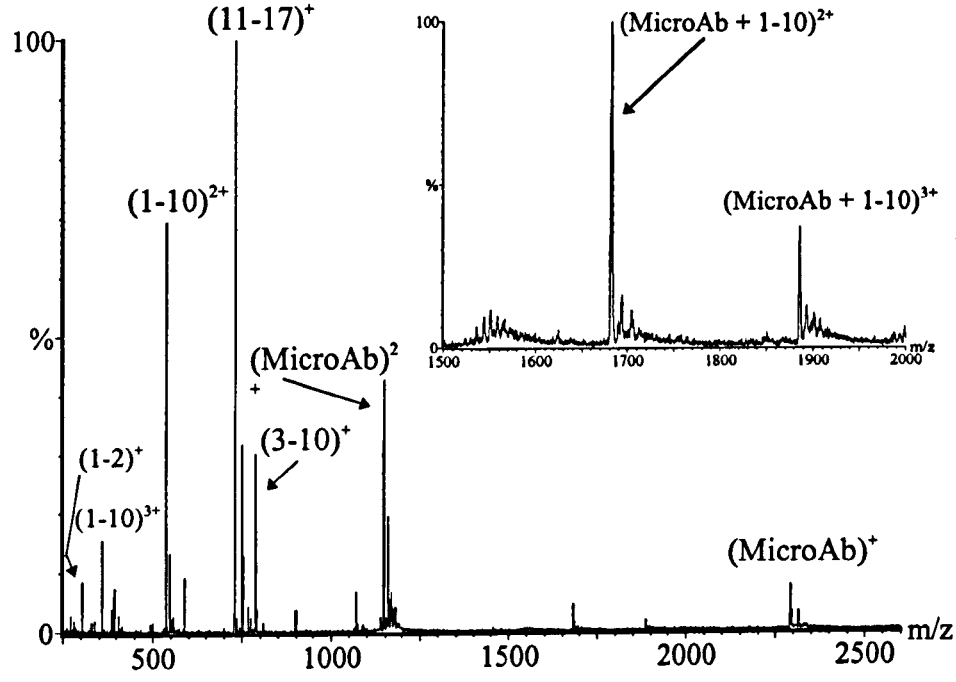


Figure 6.14. ESI spectrum of the products of a partial tryptic digest of the V3 Z-321 peptide incubated with the Micro-Ab

The Z-321 peptide was then subjected to a complete tryptic digestion the products of which are shown in Figure 6.15. As expected the most prominent peaks are due to the digest peptides containing amino acid residues 1-2, 3-10 and 11-17. The peptide containing amino acid residues 2-10 (cleavage after the arginine residue at position 1) was not produced. The products of the complete tryptic digestion of the Z-321 peptide were incubated with the Micro-Ab to determine what, if any, effect on binding to the Micro-Ab removing the RK dipeptide would have.

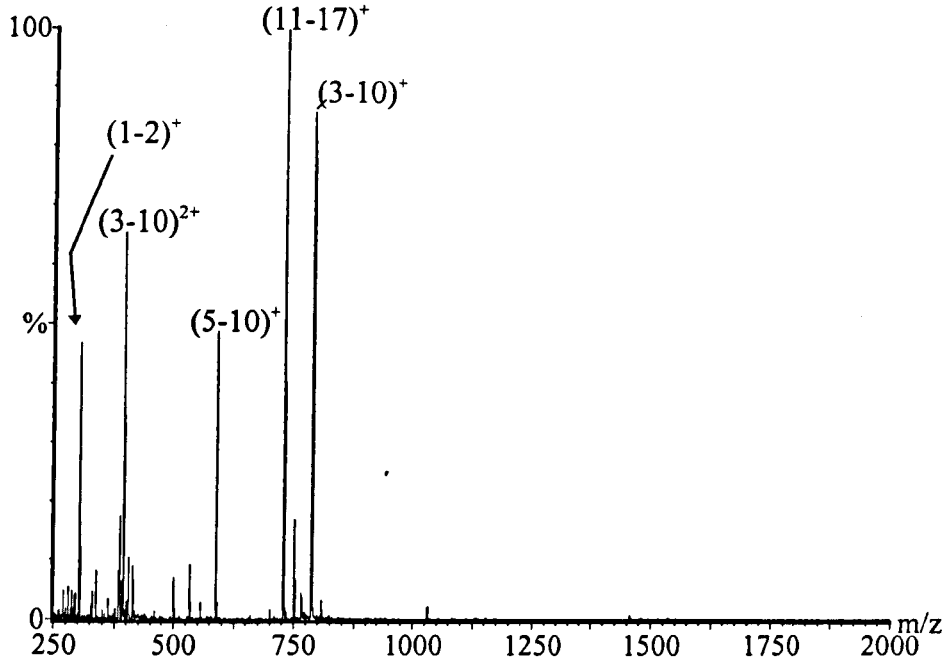


Figure 6.15. ESI spectrum of the products of a complete tryptic digestion of the V3 Z-321 peptide

Figure 6.16. shows the ESI spectrum of the incubation of the products of the complete tryptic digestion with the Micro-Ab.

The spectrum shows peaks corresponding to the digest peptides assigned previously and peaks corresponding to singly and doubly charged ions of the Micro-Ab. No binding, however, was observed between the Micro-Ab and any of the digest peptides produced by the complete tryptic digestion of the V3 Z-321. These data suggests that that peptide 1-2 (arginine-lysine) is an essential part of the Micro-Ab epitope (ie RKSISIGPGR) although it did not bind itself to the Micro-Ab.

The data from the incubation of the products of partial and complete digestions of the V3 Z-321 peptide incubated with the Micro-Ab are summarised in Figure 6.17.

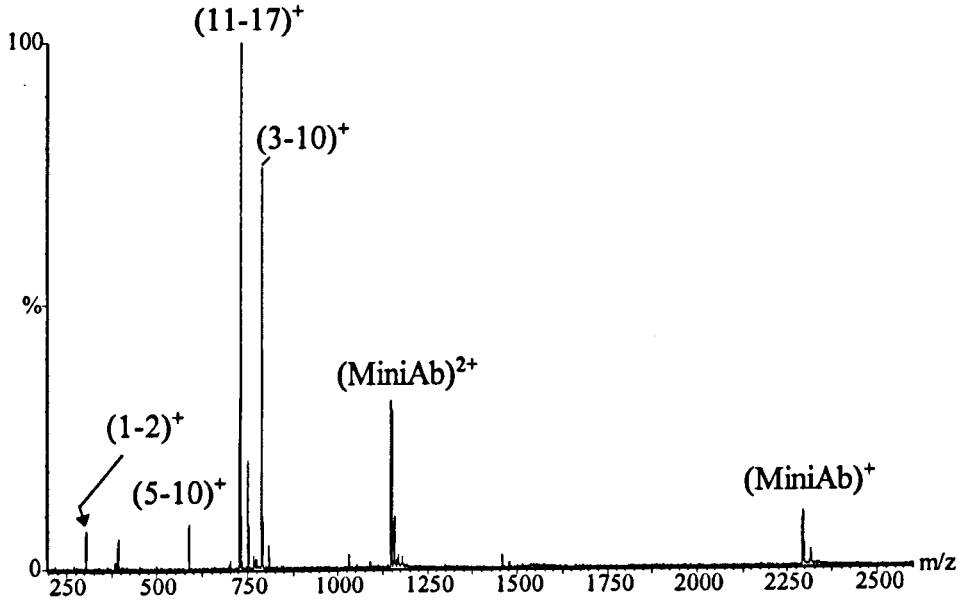


Figure 6.16. ESI spectrum of the incubation of the products of a complete tryptic digestion of the V3 Z-321 peptide incubated with the Micro-Ab

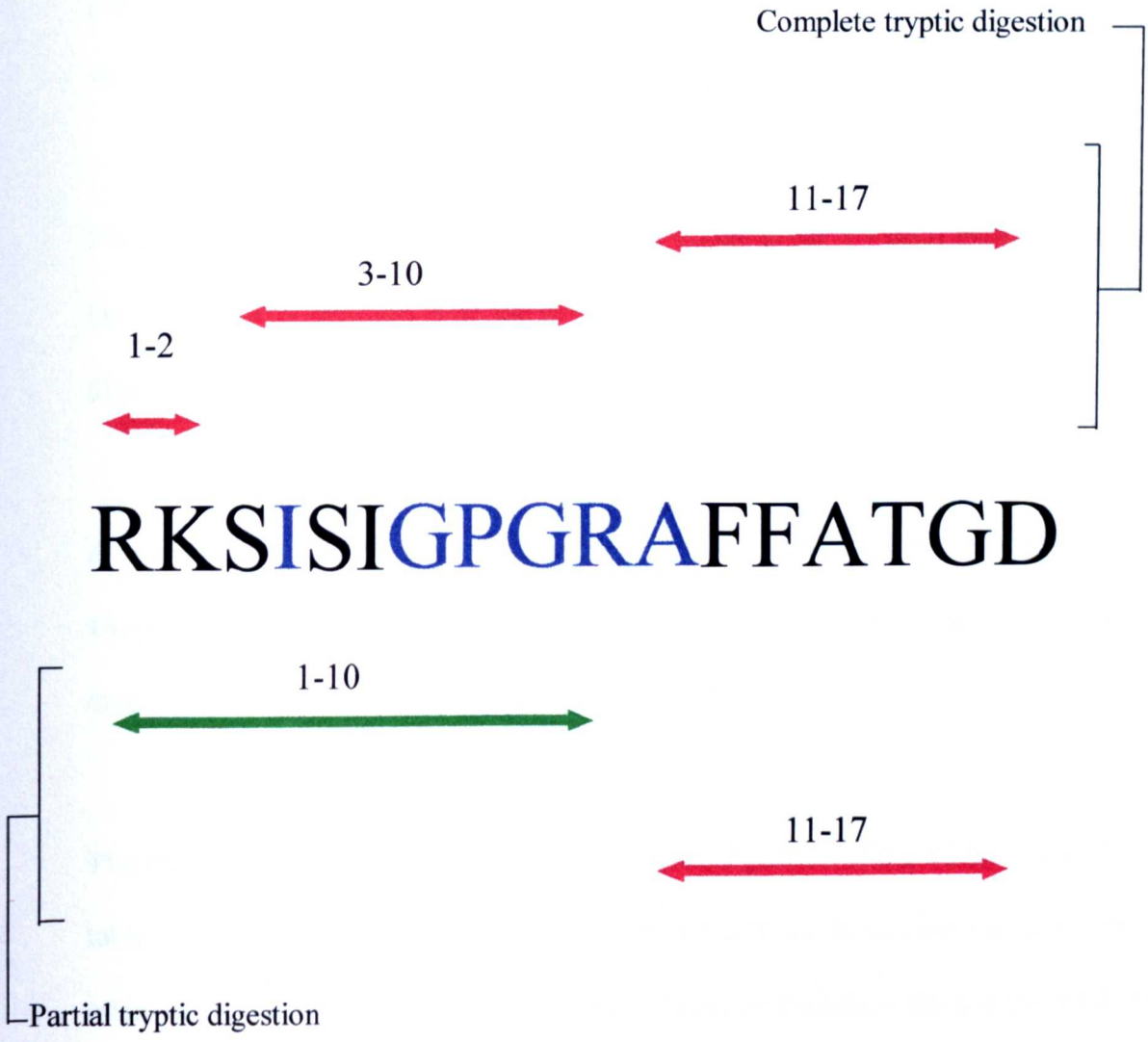


Figure 6.17. **Major peptides produced by a partial and complete tryptic digestion of the V3 Z-321 peptide. The green line represent the peptide that was bound by the Micro-Ab, the red lines represent the peptides that were not bound by the Micro-Ab.**

This result was supported by the results of a complete tryptic digest of the V3 BRA peptide. The ESI spectrum of the products of a tryptic digest of the V3 BRA peptide is shown in Figure 6.18.

No attempt to reduce and alkylate the disufide bridge in the V3 BRA peptide was made (as the epitope is removed from the terminal cysteine residues) so some of the predicted digest peptides are missing from the digest mixture.

A digest peptide was identified as being the peptide containing amino acid residues 10-18, (sequence; KSIHIGPGR). As trypsin cleaves after arginine and lysine residues this digest peptide is an unexpected tryptic digest product.

The major digest peptides produced by the trypsin digestion of the V3 BRA are listed in table 6.3. The 11-18 digest peptide omitted both the R and K residues i.e. SIHIGPGR, which were found to be necessary for binding between the Micro-Ab and the V3 Z-321 peptide, whilst the 10-18 peptide included just the K residue on the peptide i.e. KSIHIGPGR.

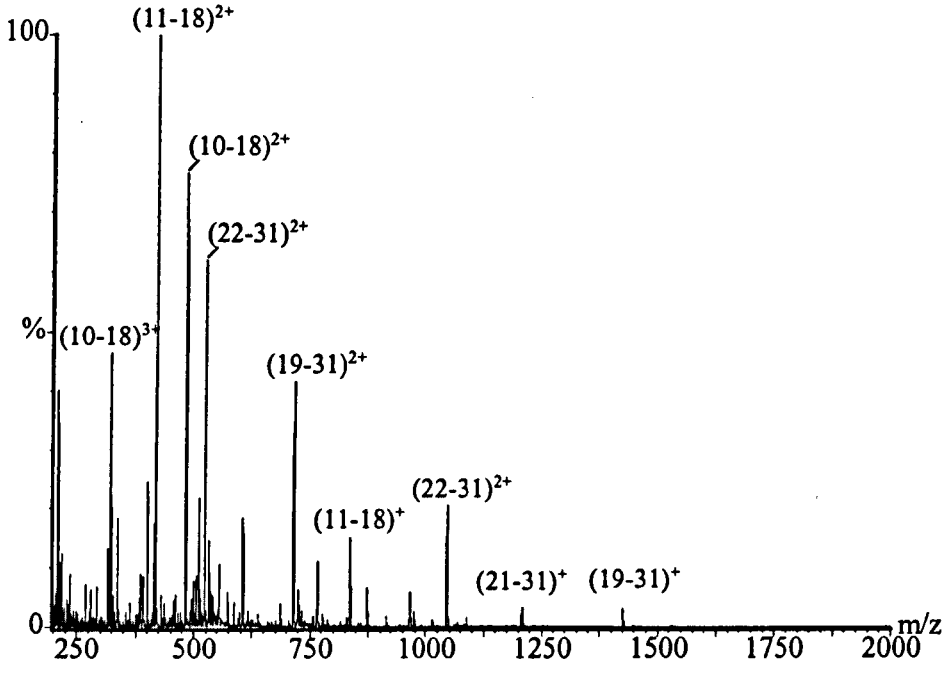


Figure 6.18. ESI spectrum of the products of a tryptic digestion of the V3 BRA peptide

BRA peptide fragment	Charge state	Measured RMM (Da)	Expected RMM (Da)	Difference (Da)
19-31	1+	1425.8	1425.7	0.1
19-31	2+	713.4	713.7	0.3
11-18	1+	836.5	836.5	0.0
11-18	2+	418.9	418.8	0.1
21-31	1+	1207.7	1207.6	0.1
21-31	2+	604.7	604.5	0.2
-1-9,32-35-	2+	765.7	765.8	0.1
-1-9,32-35-	3+	510.8	510.9	0.1
22-31	1+	1044.5	1044.5	0.0
22-31	2+	523.0	522.8	0.2
10-18	1+	964.7	964.6	0.1
10-18	2+	483.0	482.8	0.2
10-18	3+	322.7	322.7	0.0

Table 6.3. Expected and measured RMM values of the products of a tryptic digestion of the V3 BRA peptide.

Figure 6.19. shows the ESI spectrum of the products of a tryptic digestion of the V3 BRA peptide incubated with the Micro-Ab.

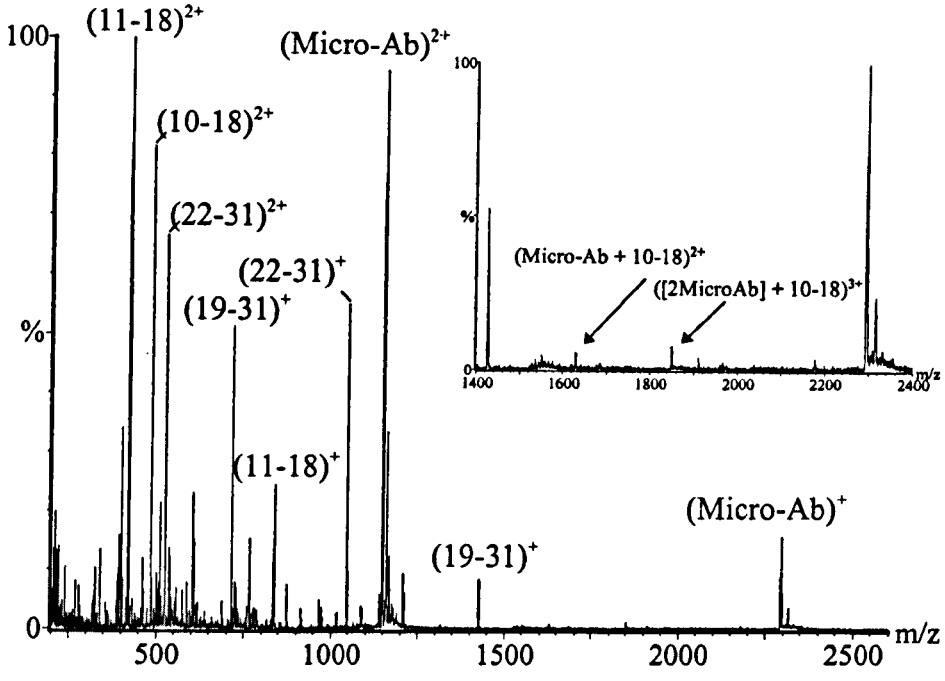


Figure 6.19. The ESI spectrum of the products of a tryptic digestion of the BRA peptide incubated with the Micro-Ab. Inset shows expanded m/z range.

The spectrum shows peaks corresponding to the digest ions as assigned previously and those due to the singly and doubly charged ions of the Micro-Ab. The inset shows peaks assigned as one and two Micro-Abs binding to the digest peptide containing amino acid residues 10-18. No binding to the digest peptide containing amino acid residues 11-8 is observed. This result suggests that the 11-18 peptide, sequence SIHIGPGR, is unable to bind but the addition of the K residue, i.e. the 10-18 peptide, is able to bind. This supports the view that both the R and K residues, or in the case of the BRA peptide just the K residue is required in addition to the SIHIGPGR sequence for binding to the

Micro-Ab to take place. The relatively small amount of binding observed betweeb the KSIHIGPGR peptide and the Micro-AB could be due to the presence of just the lysine residue as opposed to the presence of both the lysine and arginine residues as in the binding between the Micro-Ab and the Z-321 peptide.

6.4. C8029 Negative Micro-Ab binding studies

The C8029 peptide (Negative Micro-Ab) lacks one tyrosine residue from position 5 or 6 of the Micro-Ab. This peptide, in contrast to the Micro-Ab, has no activity against HIV-1. An ESI-MS study of its binding to the two HIV-1 V3 peptide was therefore undertaken to determine if the absence of the tyrosine residue would effect the binding to the V3 peptides.

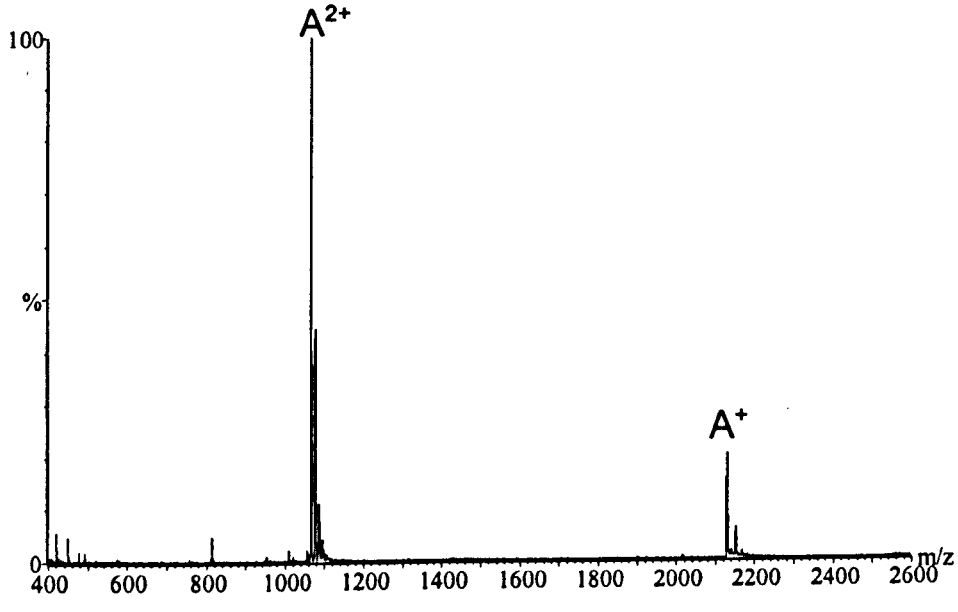


Figure 6.20. ESI spectrum of the C8029 peptide (negative Micro-Ab)

The measured RMM of the negative Micro-Ab of 2129.6 Da ±0.2 Da compares well with the calculated RMM of the peptide of 2129.8 Da (Figure 6.20). This confirms the primary sequence of the peptide and the presence of one disufide bond as in the standard Micro-Ab. The negative Micro-Ab was then incubated with the V3 peptides in a manner identical to that used for the standard Micro-Ab and the V3 peptides. Figure 6.21. shows the spectrum resulting from the incubation of the negative Micro-Ab with the V3 Z-321 peptide. It shows the typical negative Micro-Ab contributions and those due to the Z-321 peptide. The inset shows an expanded m/z range which shows peaks corresponding to triply and quadruply charged ions of a complex between the negative Micro-Ab and the V3 Z-321 peptide. This result suggests that binding between the negative Micro-Ab and the V3 Z-31 peptide can take place despite the absence of the tyrosine residue.

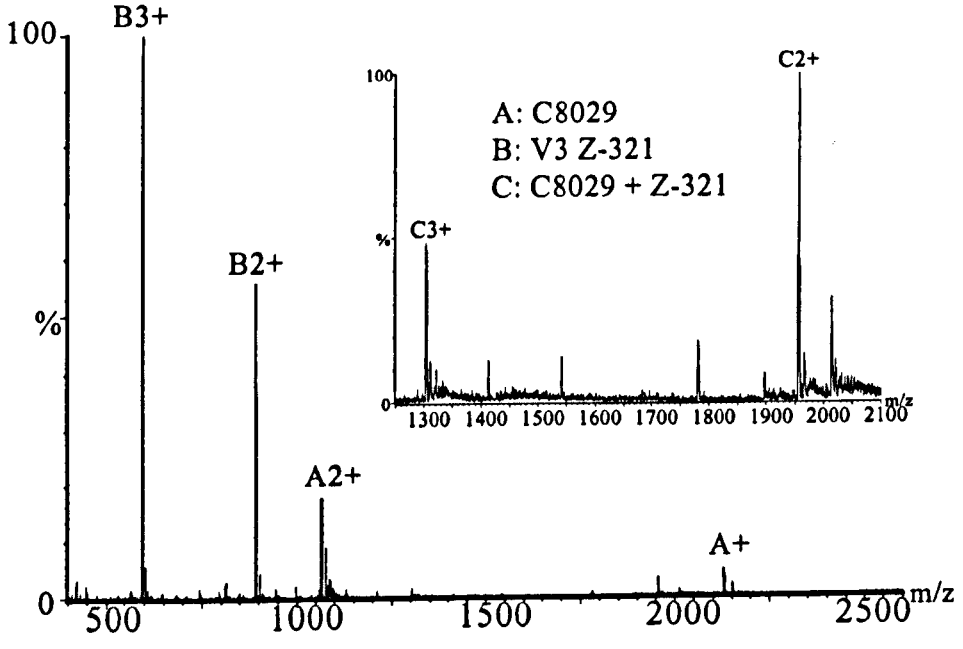


Figure 6.21. Negative Micro-Ab incubated with the V3 Z-321 peptide.

Upon incubation between the negative Micro-Ab and the V3 BRA peptide peaks corresponding to a complex between the two were also observed (spectrum not shown). The ESI spectrum of the C8029 negative Micro-Ab incubated with the products of a partial tryptic digestion of the V3 Z-321 peptide are shown in Figure 6.22. The inset shows peaks corresponding to the Negative Micro-Ab binding to the peptide containing amino acid residues 1-10 of the V3 peptide. This result indicates that the negative Micro-Ab binds in a manner identical to that of the standard Micro-Ab although, as stated previously the negative Micro-Ab shows no activity against HIV-1. Incubation of the negative Micro-Ab with the products of a complete tryptic digestion of the V3 Z-321 peptide indicates that no binding occurs to the peptide containing amino acid residues 3-10 as was found for the standard Micro-Ab (Figure 6.23.).

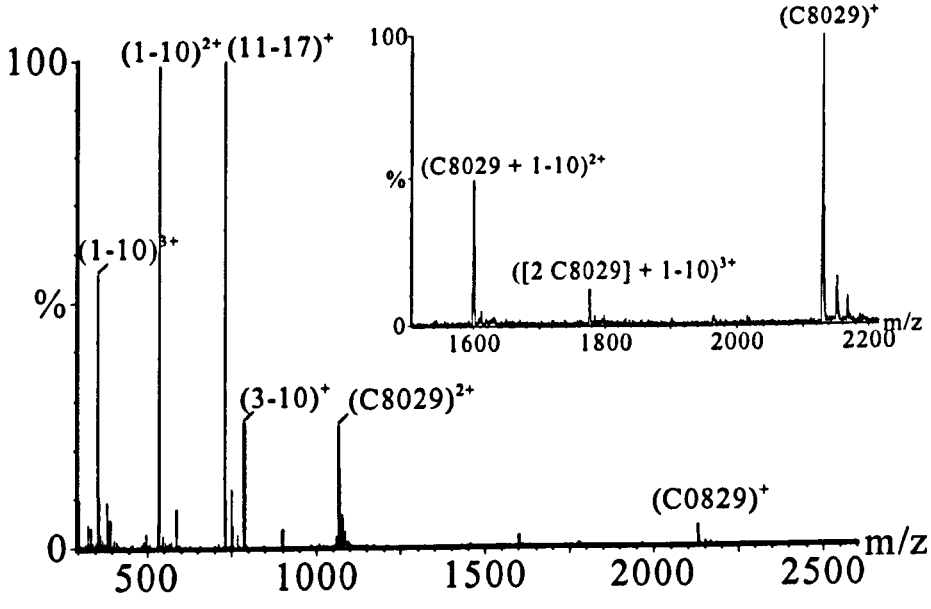


Figure 6.22. ESI spectrum of the C8029 negative Micro-Ab incubated with the products of a partial tryptic digestion of the V3 Z-321 peptide

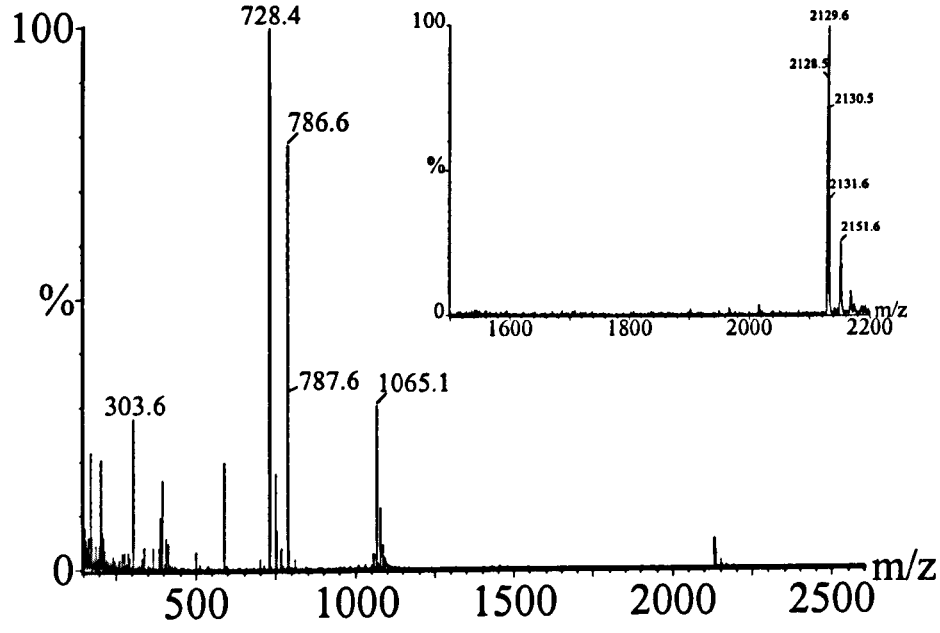


Figure 6.23. ESI spectrum of the negative Micro-Ab incubated with the products of a complete tryptic digestion of the V3 Z-321 peptide

6.5. Conclusions

The results presented in this chapter indicate that the low RMM value of the Micro-Ab, as compared to the high RMM of a typical antibody of approximately 150 kDa makes possible the investigation of its binding activity to two HIV-1 V3 peptides by means of ESI-MS analysis on a quadrupole instrument of geometry QhQ equipped with electrospray ionisation.

RMM determinations of the Micro-Ab and the V3 peptides Z-321 and BRA were within 0.1% of their calculated values. Furthermore it has been demonstrated that ESI can produce gas phase ions which correspond to a non-covalent complex between the Micro-Ab with the V3 peptides and can thus be detected by MS. The interaction between the Micro-Ab and the V3 peptides implies that its antiviral activity is mediated by binding directly to the virus particle. The ability to maintain and detect the complex between the Micro-Ab and the V3 peptides negates the need to immobilise the antibody and perform several washing procedures which characterise the majority of the MALDI-TOF-MS studies previously performed.

Partial and complete enzymatic digests of the V3 peptides produced peptides which were subsequently incubated with the Micro-Ab. The Micro-Ab selectively bound only to the peptides containing the epitope. Results indicted that the consensus epitope was RKXXXGPGRA. The dipeptide R and K was found to be required for binding to take place, although a smaller amount of binding was observed with just the K residue at the C-terminus end of the epitope in the case of the BRA peptide. Experiments to determine the epitope were completed in under two hours.

The negative Micro-Ab (lacking one tyrosine residue of the standard Micro-Ab) was also observed to bind to the V3 peptides. Incubation of the negative Micro-Ab with partial and complete digestion of the V3 peptide indicated that it binds in an identical manner to the standard Micro-Ab.

6.6. References

1. N. J. Dimmock, *Topics. Microbiol. Immunol.*, **183**, 1 (1993).

2. N. J. Dimmock, *Rev. Med. Virol.*, **5**, 165 (1995).

3. M. R. Posner, T. Hideshima, T. Cannon, M. Mukherjee, K. H. Mayer, R. A. Byrn, *J. Immunol.*, **146**, 4325 (1991).

4. T. L. McInerney, L. McLain, S. J. Armstrong, N. J. Dimmock, *Virol.*, **233**, 313 (1997).

5. S. Ugolini, I. Mondor, P. W. H. I. Parren, D. R. Burton, S. A. Tilley, P. J. Klasse, Q. J. Sattentau, *J. Exp. Med.*, **186**, 1287 (1997).

6. A. Valenzuela, J. Blanco, B. Krust, R. Franco, A. G. Hovanessian, *J. Virol.*, **71**, 8289 (1997).

7. S. J. Armstrong, N. J. Dimmock,. *J. Gen. Virol.*, **77**, 1397 (1996).

8. N. A. C. Jackson, B. Wahren, M. Levi, N. J. Dimmock. *Submitted for publication.* (1998).

9. J. A. Levy,. ASM Press, Washington, D.C. (1994).

10. D. R. Burton,. *Proc. Natl. Acad. Sci. USA.*, **94**, 10018 (1997).

11. J. A. Kessler, P. M. McKenna, E. A. Emini, C. P. Chan, M. D. Patel, S. K. Gupta, G. E. Mark, C. F. Barbas, D. R. Burton, A. J. Conley, *AIDS Res. Hum. Retroviruses,* **13**, 575 (1997).

12. M. P. D'Souza, D. Livnat, J.A. Bradac, S.H. Bridges,. *J. Infect. Dis.*, **175**, 1056 (1997).

13. T. Muster, F. Steindl, M. Purtscher, A. Trkola, A. Klima, F. Ruker, G. Himmler, H. Katinger., *J. Virol.*, **67**, 6642 (1993).

14. D. Suckau, J. Kohl, G. Karwath, K. Schneider, M. Casaretto, D. Bitter-Suermann, M. Przybylski, *Proc. Natl. Acad. Sci. USA,* **87**, 9848 (1990).

15. D. I. Papac, J. Hoyes, K. B. Tomer, *Anal. Chem.*, **66**, 2609 (1994).

16. C. E. Parker, D .I. Papac, S. K. Trojak, K. B. Tomer, *J. Immunolog,* **157**, 198 (1996).

17. Y. Zhao, B.T. Chait, Anal. Chem., **66**, 3723 (1994).

18. S. Jeyarajah, C. E. Parker, M. T. Sumner, K. B. Tomer,. *J. Am. Soc. Mass Spectrom,*. **9**, 157 (1998).

19. M. Macht, W. Fiedler, K. Kurzinger, M. Przybylski, *Biochemistry,* **35**, 15633 (1996).

20. Y. V. Lyubarskaya, Y. M. Dunayevskiy, P. Vouros, B. L. Karger, *Anal. Chem.*, **69**, 3008 (1997).

21. Y. Lu, S. J. Gaskell, J. L. Brookman.. *J. Am. Soc. Mass Spectrom.*, **9**, 208 (1998).

22. M. Levi, M Sallberg, U. Ruden, D. Herlyn, H. Maruyama, H. Wigzell, J. Marks, B. Wahren. *Proc. Natl. Acad. Sci. USA.*, **90**, 4374 (1993).

23. L. Akerblom, J. Hinkula, P. A. Broliden, B. Makitalo, T. Fridberger, J. Rosen, M. Villacres-Erikssen, B. Morein and B. Wahren,. *AIDS,* **4**, 953 (1990).

24. P. A. Broliden, B. Makitalo, L. Akerblom, J. Rosen, K. Broliden, G. Utter, M. Jondal, E. Norrby, B. Wahren, *Immunol.*, **73**, 371 (1991).

25. P. A. Broliden, K. Ljunggren, J. Hinkula, E. Norrby, L. Akerblom, B. Wahren, *J. Virol.*, **64**, 936 (1990).

26. H. R. Gelderblom, *AIDS*, **5**, 617 (1991).

27. W. Weissenhorn, S. A. Wharton, L. J. Calder, P. L. Earl, B. Moss, E. Aliprandis, J. J. Skehel, D. C. Wiley, *EMBO J.*, **15**, 1507 (1996).

28. D. C. Chan, D. C. Fass, J. M. Berger, P. S. Kim, *Cell*, **89**, 263 (1997).

29. W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel, D. C. Wiley, *Nature*, **371**, 37 (1997).